

```

Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 1861 GGCATGCCAGCCTTGCCCATCACTGCACACCACTCTCTGTGGACCTGGATGACAAG 1920
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCGAGCAGAGAGCCCTCTGACGTCCTATCTCTGCGGTGGTGGC 1980
Qy 653 653
Db 1981 ATTCTGCTGGTCTGGTCTTGGGGTGTCTTTGGGATCCTTCATCAGCGACGCCAGCAG 2040
Qy 653 653
Db 2041 AAGATCCGGAAGTACACGATCGGAGACTGCTGCAGGAACGGAGCTGGTGGAGCCGCTG 2100
Qy 653 653
Db 2101 ACACCTAGCGAGCGATGCCAACCCAGCGCAGATCGGATCTCTGAAAGACAGCGAGCTG 2160
Qy 653 653
Db 2161 AGGAAGTGAAGTCTGGATCTGGCGCTTTTGGCACAGTCTACAGGGCATCTGGATC 2220
Qy 653 653
Db 2221 CCTGATGGGAGATGTGAAATTCAGTGCCCATCAAAAGTGTGAGGGAAACACATCC 2280
Qy 653 653
Db 2281 CCCAAGCCCAACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGTGGGTGCCCA 2340
Qy 653 653
Db 2341 TATGTCTCCGCTCTCTGGCATCTGCTGACATCCACGCTGACAGCTT 2400
Qy 653 653
Db 2401 ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCCGCGACGCTTGGGCTCCAG 2460
Qy 653 653
Db 2461 GACCTGCTGAATGCTATCCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTCGG 2520
Qy 653 653
Db 2521 CTGTTACACAGGACTTGGCGGCTCGGAAGCTGTGTCAGAGTCCCAACCATGTCAA 2580
Qy 653 653
Db 2581 ATTACAGACTTGGGCTGGCTCGGCTGCTGGACATTCAGACATGACGACAGAGTACCATGCAGAT 2640
Qy 653 653
Db 2641 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTTCTCCCGCGGCTTACC 2700
Qy 653 653
Db 2701 CACCAGATGATGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2760
Qy 653 653
Db 2761 AAACCTTACGATGGATCCAGCCCGGGAGATCCCTGACCTGCTGGAAAGGGGAGCGG 2820
Qy 653 653
Db 2821 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATGATGTTGATG 2880

```

```

653 ----- 653
2881 ATTCACTCTGAATGTCGGCCAAGATTCCGGAGTTGGTGTCTGAATTTCTCCGCATGCCC 2940
654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
2941 AGGACCCCCAGCGCTTTGTGTCATCAGAAATGAGGACTTTGGGCCAGCCAGTCCCTTG 3000
665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684
3001 GACAGCACTTCTACCGCTCTACTGCTGAGGACCATGACATGGGGGAGCTGGTGGATGCT 3060
685 GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly 704
3061 GAGGAGTATCTGGTACCCAGCAGCGCTTCTCTTCACAGACCTGCCCGCGCGCTGG 3120
705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
3121 GGCATGCTCCACACAGCACCGCAGCTCTCTACAGGAGTGGCGGTGGGAGCTTGACA 3180
725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
3181 CTAGGCTGGAGCCCTCTGAAGAGAGGCCCCAGGTCTCCACTGGCACCTCCGAGGG 3240
745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
3241 GCTGGCTCCGATGTTTGTGATGGTGACCTGGGAATGGGGCAGCAAGGGGCTGCAAGC 3300
765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
3301 CTCCCCACATGACCCAGCCCTCTACAGCGTACAGTGGAGGAGCCACACAGTACCCCTG 3360
785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
3361 CCCTCTGAGACTGATGCTAGCTTGGCCCCCTGACCTGACGCCCCAGCTGGAATATGTG 3420
805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
3421 AACCAAGCAGATGTCGGCCCCAGCCCTTCGCCCGAGAGGGGCTCTGCCTGCTGCC 3480
825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
3481 CGACCTGCTGGTGCACTCTGGAAAGGCCCAAGACTCTCTCCCGAGGAGAAATGGGGTC 3540
845 ValLysAspValPheAlaPheGlyValAlaValGluAsnProGluTyrLeuThrProGln 864
3541 GTCAAGACCTTTTGGCTTTGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCAG 3600
865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
3601 GGAGGAGCTGCCCTCAGCCCCACCTCTCTGCTTCCAGCCAGCTTCGACCAACCTC 3660
885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
3661 TATTACTGGAGCAGGACCCAGAGCGGGGGCTCCACCCAGACCTTCAAGAGGGACA 3720
905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
3721 CCTACGACAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG 3765

```

RESULT 2

US-09-930-125-1

```

; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS

```

```

; TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
; US-09-930-125-1

```

Alignment Scores:

```

Pred. No.: 0 Length: 3768
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 10 Gaps: 1

```

SEQ6 (1-919) x US-09-930-125-1 (1-3768)

```

QY 1 MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProProGlyAla 20
DB 1 ATGGAGCTGGCGCCTTGTGCGCTGGGGCTCTCTCGCCCTCTTGCCTCCCGCGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
DB 61 GCGAGCACCAAGTGTGACCGGCACAGACATGAAGTCGGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
DB 121 ACCACCTGGACATGCTCCGCCACCTCTACAGAGGCTGCCAGGTGTGCGAGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
DB 181 GAACTCACCTACTGCCCCACCAATGCCAGCTGTCTCTCCTGCAGGATATCCAGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
DB 241 CAGGGCTAGCTGCTCATCGCTCACAAACAGTGAAGGAGGCTCCCTGCAGAGGTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
DB 301 ATTGCGGAGGACCCAGCTCTTTGAGGACACTATGCCCTGGCGGTGTAGACATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
DB 361 GACCCGCTGAACAATACCACTGTCTACAGGGGCTCCCGAGGAGGCTGCGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGlnIleLeuLysGlyValLeuLeuGlnArgAsnProGln 160
DB 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTATCCAGCGAACCCCGAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
DB 481 CTCCTGCTACGAGACACGATTTTGTGGAAGGACATCTTCCACAAGAACACAGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAA 600
QY 201 GlySerArgCysTrpGlyGluSerGluAspCysGlnSerLeuThrArgThrValCys 220
DB 601 GCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTGTACAGGCTCAGCGCTGCTGCTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCys 240
DB 661 GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCCATGAGCAGTGT 720

```

```

QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 721 GCTGCGGCTGCAGGGGCCCAAGCACTGTACTGCTGCGCTGCGCTGCACTTCAACACC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
DB 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGGTCACTACAACACACACAGCTTTGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
DB 841 TCCATGCCCAATCCCGAGGGCCGTATACATTCGGCGCCAGCTGTGTGACTGCTGTCCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
DB 901 TACAACTACCTTTCTACGACCTGGGATCCTGCACCTGCTGCTGCCCCCTGCACAACCAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGACGAAGCCCTGTGCCCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1021 GTGTCTATGGTCTGGGCATGGACACTTGCAGAGGTGAGGCGACTTACCAGTGCCT 1080
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
DB 1081 ATCCAGGAGTTTGTGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCGGAGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
DB 1141 TTTGATGGGGAGCCAGCCCTCCAACTGCGCCGCTCCAGCCAGAGCAGTCCAAGTGT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
DB 1201 GAGACTCTGGAGAGATCACAGTTACCTATACATCTCAGCATGGCCGCGCAGCCTGCCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
DB 1261 GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACAATGGCGC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
DB 1321 TACTGCTCAGCCTGCAGGGCTGGGCATCAGCTGGCTGGGCTGGCTGCCTCAGTGGAGAA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
DB 1381 CTGGGCAGTGGACTGCCCTCATCCACCATACACCCACCTCTGCTTGTGTCACACGGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
DB 1441 CCCTGGGACAGCTCTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
DB 1501 GAGGACGAGTGTGTGGCGAGGCGCTGGCTGCCACAGCTGTGCGCCCGAGGCACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
DB 1561 TGGGGTCCAGGGCCCAACCCAGTGTGTCAACTGCAGCCAGTCTCTTCGGGGCCAGGAGTG 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
DB 1621 GTGGAGGAATGCGAGTACTGAGGGGCTTCCCGCAGGAGTATGTGAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
DB 1681 TTGCGCGTGCACCCCTGAGTGTGAGGAGTGTGAGGAGTGTGAGTGTGAGTGTGAGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
DB 1741 GCTGACCACTGTGTGGCCTGTGCCACTATAAGGACCCCTCTTCTGCTGCGTGGCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProaspGluGlu 620

```



```
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (3765)
US-10-313-644-1
```

Alignment Scores:

```
Pred. No.: 0 Length: 3768
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 12 Gaps: 1
```

```
SEQ6 (1-919) x US-10-313-644-1 (1-3768)
```

```
QY 1 MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProProGlyAla 20
DB 1 ATGGAGCTGGCGCCTTGGCGCTGGGGCTCCTCGCCCTCTTGGCCCGCGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
DB 61 GCGAGCACCAAGTGTGACCGGCACAGACATGAAGCTGCGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
DB 121 ACCCAGCTGGACATGCTCGCCACCTCTACAGAGGCTGCCAGGTGGTCAGGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
DB 181 GAACTCACCTACCTGCCACCAATGCCAGCTGTCTCTCCAGGATATCCAGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
DB 241 CAGGGCTAGTGCTCATGCTCATCAACCAAGTGAGCGAGGTCCCACTGCAGAGGTGGCG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
DB 301 ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTCTAGCAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
DB 361 GACCCGCTGAACAATACCAACCCCTGTTCACAGGGGCTCCCCAGAGGCTGCGGGAGTG 420
QY 141 GlnLeuArgSerLeuThrGlnIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
DB 421 CAGCTTCGAGGCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
DB 481 CTCCTGCTACCGAGCACAGATTTGTGGAAGGACATCTCCACAAGAACCAACAGTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
DB 541 CTCACACTGATAGACCAACACGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAA 600
QY 201 GlySerArgCysTrpGlyGluSerGluAspCysGlnSerLeuThrArgThrValCys 220
DB 601 GGTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTACAGGCTCAGCGCCTGCTGTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys 240
DB 661 GCCGCTGGCTGTCGCCGTGCAAGGGCCACTGCCCACTGACTGTGCTCAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 721 GCTGCCGCTGTCAGCGGGCCCCAAGCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
```

```
DB 781 AGTGGCATCTGTGAGCTGCACCTGCCAGGCCCTGGTCACTACACACAGACAGCTTTGTAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
DB 841 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGGGCCAGCTGTGTACTGCTGCTGCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
DB 901 TACAACCTACCTTCTACGGACCTGGGATCTGCACCTCTGCTGCCCTGCACAAACAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 961 GAGGTGCACAGCAGAGATGGAACACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1021 GTGTCTATGTCTGGGCATGGAGCCTTGCAGAGGTGAGGGCAGTTACCAAGTGCCTAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
DB 1081 ATCCAGGAGTTTGTCTGGCTGCAAGAAGATCTTTTGGAGCTTGGCATTTCTGCCGGAGAC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
DB 1141 TTTGATGGGGACCCAGCCTCCAACTGCCCCGCTCCAGCCAGAGCAGCTCCAAAGTTT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
DB 1201 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCCGAGACGCTGCCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
DB 1261 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGACGAATCTGCACAATGGCGC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
DB 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTAGGGAA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
DB 1381 CTGGCAGTGGACTGSCCTCATCCACCATAACACCCACCTCTGCTTCTGTCACACGGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
DB 1441 CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTGTCTCCACTGCCAACCGGCCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
DB 1501 GAGGACGAGTGTGTGGCGAGGGCTTGGCTGCCACCCAGCTGTGCCCGAGGGCAGCTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
DB 1561 TGGGCTCCAGGGCCCAACCCAGTGTGTCACTGCAGCCAGTTCTTTCGGGGCCAGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
DB 1621 GTGGAGGAATCCGAGTACTGCGGGGCTCCCGAGGGAGTATGTAAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
DB 1681 TTGCCGCTGCCACCCCTGAGTGTGAGCCCAAGATGCTCAGTACCTGTTTGTGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
DB 1741 GCTGACCACTGTGGCTGTGCCCTGTGCCACATATAAGGACCTCCCTTCTGCTGCGCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
DB 1801 CCCAGCGGTGTGAAACCTGACCTCTCATACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
DB 1861 GCGCATGCCAGCTTGGCCCATCAACTGCCCCACTCTCTGTGTGACCTTGTATGACAAG 1920
```

```
Qy 641 GlyCysProAlaGluInArgAlaSerProLeuThrSer----- 653
Db 1921 GGCTGCCCGCGAGCAGAGAGCCAGCCCTCTGAGCTCCATCATCTCTGCGGTGGTTGGC 1980
Qy 653 ----- 653
Db 1981 ATTCTGCTGGTGGTGGTCTTGGGGTGGTCTTTGGGATCCCTCATCAAGCGACGCGAGCAG 2040
Qy 653 ----- 653
Db 2041 AAGATCCGGAAGTACACGATCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCGCGGTG 2100
Qy 653 ----- 653
Db 2101 ACACCTAGCGGAGCGATGCCCAACACGCGCAGATGCGGATCCTGAAAGACGCGAGGTG 2160
Qy 653 ----- 653
Db 2161 AGGAAGGTGAAGTCTTGGATCTGGCCTTTTGGCACAGTCTTACAAAGGGCATCTGGATC 2220
Qy 653 ----- 653
Db 2221 CCTGATGGGAGAAATGGAATAATCCAGTGCCCATCAAAAGTGTGAGGGAAAAACATACC 2280
Qy 653 ----- 653
Db 2281 CCCAAAGCAACAAGAAATCTTAGACGAGCATACGTGATGGCTGTGGTGGCTCCCA 2340
Qy 653 ----- 653
Db 2341 TATGTCTCCCGCTTCTGGGCATCTGCTGACATCCACGGTGACGTGGTGACACAGCTT 2400
Qy 653 ----- 653
Db 2401 ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCGCGAGCGCTGGGTCCAG 2460
Qy 653 ----- 653
Db 2461 GACCTGTGNACTGGTGTATGCAGATTCCCAAGGGGATGACCTACCTGGAGATGTGGG 2520
Qy 653 ----- 653
Db 2521 CTCTACACAGGACTTGGCGCTCGGAACGTGCTGTCAAGAGTCCCAACCATGTCAAA 2580
Qy 653 ----- 653
Db 2581 ATTACAGACTTGGGCTGGCTCGGCTGCTGGACATTCACGAGACAGAGTACCATGCAGAT 2640
Qy 653 ----- 653
Db 2641 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGTTTACC 2700
Qy 653 ----- 653
Db 2701 CACCAGATGATGTGTGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGCGC 2760
Qy 653 ----- 653
Db 2761 AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGTGTGAAAGGGGAGCGG 2820
Qy 653 ----- 653
Db 2821 CTGCCCCAGCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAATGTGGATG 2880
Qy 653 ----- 653
Db 2881 ATTGACTCTGAATCTCGGCCAAGATTCGGGAGTTGGTGTCTGAATTCCTCCGCGATGGCC 2940
Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 2941 AGGACCCCCAGCGCTTTGTGGTCAATCAGAAATGAGGACTTGGGCCCGCAGCGCTCCCTTG 3000
```

```
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACTTCTACCGCTCACTGCTGGAGGACATGACATGGGGAGCTGGTGGATGCT 3060
Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGGTACCCAGCAGGGCTTCTTGTCTCAGACCTTGCCTCCGCGCTGGG 3120
Qy 705 GlyMetValHisAlaGlnHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
Db 3121 GGCATGTCCACACAGCGACCGAGCTCATCTACAGGAGTGGCGGTGGGAGCTGACA 3180
Qy 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3181 CTAGGGCTGGAGCCCTCTGAAGAGAGAGGCCCCAGGTCTCCACTGGCACCTCCGAAGG 3240
Qy 745 AlaglySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3241 GCTGGCTCCGATGTATTTGATGGTACCTGGGAATGGGGCAGCCAAAGGGCTGCAAGC 3300
Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3301 CTCCCCACATGACCCCGCCCTCTACAGCGGTACAGTGAGGAGCCCGACAGTACCCCTG 3360
Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3361 CCCCTGAGACTGATGGCTACGTTGCCCTGACCTGCAGCCCCAGCTGAATATGTG 3420
Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3421 AACAGCCAGATGTTCGGCCCGCCCTTCGCCCGCAGAGAGGCCCTCTCCCTGCTGCC 3480
Qy 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3481 CGACTCTGCTGGCCACTCTGGAAGGCCCAAGACTCTCTCCCCAGGGAAGATGGGGTC 3540
Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3541 GTCAAGACGCTTTTGGCTTTGGGGTGGCTGGAGAACCCCGAGTACTTGACACCCAG 3600
Qy 865 GlyGlyAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3601 GGAGGAGCTGCCCTCAGCCCCACCTCTCTGCTTACGCCAGCGCTTCGACAACTC 3660
Qy 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3661 TATTACTGGGACCAAGGAGCCACCAGAGCGGGGGCTCCACCCAGACCTTCAAAGGGACA 3720
Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3721 CCTACGGCAGAGAACCCAGAGTACTGGGTCTGGAGCTGCCAGTG 3765
```

RESULT 4

```
US-09-441-411-5
; Sequence 5, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441.411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-5
```

Alignment Scores:

Pred. No.:	0	Length:	4473
Score:	4900.00	Matches:	919
Percent Similarity:	73.23%	Conservative:	0
Best Local Similarity:	73.23%	Mismatches:	0
Query Match:	96.49%	Indels:	336
DB:	11	Gaps:	1

SEQ6 (1-919) x US-09-441-411-5 (1-4473)

QY	1	MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla	20
DB	175	ATGGAGCTGGCGGCTTGGCGCTGGGGCTGCCCTCGCCCTCTTGGCCCGCGAGCC	234
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
DB	235	GCAGACCAAGTGTGCACCGGCACAGACATGAAGCTGCGCTCCCTGCCAGTCCCGAG	294
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
DB	295	ACCCACCTGGACATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGGTGAGGAACCTG	354
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
DB	355	GAACCTACCTACTGCCCCACCAATGCCACCTGTCTTCTCGAGGATATCCAGGAGTG	414
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
DB	415	CAGGGCTACGTCTCATCGCTCACAAACCAAGTGAGGAGTCCCACTGCAGAGGTGCGG	474
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
DB	475	ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCGCGTGTAGACAAATGA	534
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
DB	535	GAACCGCTGAACATACCAACCCCTGTACAGGGGCTCCCCAGGAGGCTCGGGAGCTG	594
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
DB	595	CAGCTTCGAAGCTCACAGAGATCTTGAAGAGGGGTCTTGATCAGCGGAACCCCCAG	654
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	180
DB	655	CTCTGTCTACGAGACAGATTGTGGAAGGACATCTTCCACAAGAACCAACAGTGGCT	714
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
DB	715	CTCACACTATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAG	774
QY	201	GlySerArgCysTrpGlyGluSerGluAspCysGlnSerLeuThrArgThrValCys	220
DB	775	GGCTCCCGCTGTGGGAGAGATTCTGAGGATTGTACAGAGCTGACGCGCACTGTCTGT	834
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCys	240
DB	835	GCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCCATGAGCAGTG	894
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
DB	895	GCTGCGGGTGCACGGGCCCCAAGCACTCTGACTGGCTGGCTGCCCTCCCACTTCAACCA	954
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
DB	955	AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACACACACACACAGTTTGA	1014
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
DB	1015	TCCATGCCCAATCCCGAGGCGGGTATACATTGCGCGCCAGCTGTGTGACTGCCGTGCC	1074
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
DB	1075	TACAACTACTTTCTACGAGACGTGGGATCTGCACCTCGTCTGCCCCCTGCACACCAA	1134

QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
DB	1135	GAGGTGACAGCAGAGAGTGAACACACAGCGGTGTGAGAAGTGACGAAGCCCTGTGCCCA	1194
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
DB	1195	GTGTGCTATGTGTGGCATGGAGCACTTGCAGAGGTTGAGGCACTTACCAGTGCCTA	1254
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
DB	1255	ATCCAGGAGTTGTGTGGTGCAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC	1314
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGlnLeuGlnValPhe	400
DB	1315	TTTGTATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGAGCAGTCCAAGTGT	1374
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
DB	1375	GAGACTCTGGAAGAGATCACAGTTACCTATATACATCTCAGCATGGCCGACAGCTGC	1434
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
DB	1435	GACCTCAGCGCTCTCCAGAACCTGCAAGTAAATCCGGGAGCAATCTCTGCACAAATGG	1494
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
DB	1495	TACTCGTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTCGCTCACTAGGGAA	1554
QY	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
DB	1555	CTGGCAGTGGACTGGCCCTCATCCACCATAACACCCACTCTCTGCTGTCGACACGGT	1614
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
DB	1615	CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTCTCCACTGCCAACCGGCCA	1674
QY	501	GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
DB	1675	GAGGACGAGTGTGTGGCGAGGCGCTGGCTGCCACAGCTGTGCCCGCGAGGCACTGC	1734
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
DB	1735	TGGGTGCCAGGCGCCACCCAGTGTGTCAACTGCAGCAGTCTCTCGGGCGCAGGAGTG	1794
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
DB	1795	GTGGAGGAATGCCGAGTACTGCAGGGCTCCCGAGGAGTATGTGAATGCCAGCACTGT	1854
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
DB	1855	TTGCCGTGCCACCTGAGTGTAGCCCAAGATGGCTCAGTACCTGTGTGGACCGGAG	1914
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
DB	1915	GCTGACCACTGTGTGGCTGTGCCCACTATAAGGACCTCCCTCTCTGCGGCGCGCTGC	1974
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
DB	1975	CCCAGGGTGTGAACCTGACCTCTCTACATGCCCATGCCCATCTGGAAGTTTCCAGATG	2034
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
DB	2035	GGCGCATGCCAGCTTGCCCCCATCACTGCACCCCACTCCCTGTGTGGACCTGGATGACA	2094
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
DB	2095	GGCTGCCCGCGAGCAGAGCCGCTCTGACGTCCATCATCTCTCGGTGTGTGC	2154
QY	653	-----	653
DB	2155	ATTCTGCTGGTGTGTGTGGGGGTGTCTTTGGGATCCTCATCAAGCGACGCGACGAG	2214

```
QY 653 ----- 653
Db 2215 AAGATCCGGAAGTACACGATCGGAGACTGTCAGGAAACGAGCTGGTGAGCCGCTG 2274
QY 653 ----- 653
Db 2275 ACACCTAGCGAGCGATGCCAACACGCGCAGATGCGGATCCTGAAAGACGAGCTG 2334
QY 653 ----- 653
Db 2335 AGGAAGTGAAGTCTTGGATCTGGCCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2394
QY 653 ----- 653
Db 2395 CCTGATGGGAGAAATGTGAAAATTCACGTGGCCATCAAAGTGTGAGGGAAAAACATCC 2454
QY 653 ----- 653
Db 2455 CCCAAGCCAAAGAANAATCTTAGACGAAGCATACGTGATGGTGTGGGCTCCCA 2514
QY 653 ----- 653
Db 2515 TATGCTCCGCTTCTGGGATCTGCCTGACATCCACGGTGCAGCTGGTCACACAGCTT 2574
QY 653 ----- 653
Db 2575 ATGCCCTATGGCTGCTCTTAGACCATGTCGGGAAACCGCGACGCTGGGCTCCAG 2634
QY 653 ----- 653
Db 2635 GACCTGTGAATCGGTATGACAGATTGCCAAGGGGATGAGCTACCTGGAGATGTGCGG 2694
QY 653 ----- 653
Db 2695 CTCGTACACAGGACTTGGCGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAAA 2754
QY 653 ----- 653
Db 2755 ATTACAGACTTGGGCTGGCTCGGCTGTGACATTGACGAGACAGATACCATGCAGAT 2814
QY 653 ----- 653
Db 2815 GGGGGCAAGTGGCCCAAGTGGATGGCGTGGAGTCCATTTCGCCCGCGGCTTCACC 2874
QY 653 ----- 653
Db 2875 CACCAGATGATGTGTGAGTTATGCTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2934
QY 653 ----- 653
Db 2935 AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGGGCGG 2994
QY 653 ----- 653
Db 2995 CTGCCCCAGCCCCCTCTGCACCATTTGATGCTACATGATCATGTTCAAATGTTGATG 3054
QY 653 ----- 653
Db 3055 ATTGACTCTGAATGTCGGCAAGATTCCGGGAGTTGGTGTGTAATTCCTCCCGCATGGCC 3114
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 3115 AGGACCCCCAGCGCTTTGTGGTTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTG 3174
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspMetGlyAspLeuValAspAla 684
Db 3175 GACAGCACCTTCTACCGCTCCTGCTGGAGACCATGACATGGGGGACCTGGTGGATGCT 3234
QY 685 GluGlyTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3235 GAGGAGTATCTGGTACCCAGCAGGGCTTCTTGTCTCCACACCTGCCCCGGGGCTGGG 3294
QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
```

```
Db 3295 GGCATGTCTCCACACAGGCACCGAGCTCACTTACCAGGAGTGGCGTGGGAGCTGCACA 3354
QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3355 CTAGGCTGGAGCCCTCTGAAGAGAGGCCCCAGGTCTCCACGTGGCACCTCCGAAGGG 3414
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3415 GCTGGCTCCGATGTATTTGATGTGACCTGGGAATGGGGCAGCAAGGGGCTGCAAGC 3474
QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3475 CTCCCCACATGATGACCCAGCCCTCTACAGCGGTACAGTGAAGAGCCACACAGTACCCCTG 3534
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3535 CCCTCTGAGACTGATGGCTAGCTTGGCCCCCTGACCTGCAGCCCCAGCCCTGAATATGTG 3594
QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3595 AACCCAGCATGTTCGGCCCCAGCCCTTCGCCCGGAGAGGGCCCTCTGCTGCTGCC 3654
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3655 CGACCTGCTGGTGCCTCTGGAAGGCCCAAGACTCTCTCCCGAGGGAAGATGGGTC 3714
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3715 GTCAAGACGCTTTTGGCTTGGGGTGGCGGTGGAGAACCCCGAGTACTTGGACACCCAG 3774
QY 865 GlyGlyAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3775 GGAGGAGTGGCTCAGCCCCACCTCTCTGCTGCTTACGCCAGCCCTTCGACACACCTC 3834
QY 885 TyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3835 TATTACTGGGACCAAGGACCACCAAGCGGGGCTCCACCCAGCACCTTCAAAGGACA 3894
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3895 CCTACGGCAGAAACCCAGAGTACCTGGGTCTGGACGTGGCAGTG 3939

RESULT 5
US-10-101-510-81
; Sequence 81, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117,0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-81

Alignment Scores:
Pred. No.: 0 Length: 4473
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 12 Gaps: 1

SEQ6 (1-919) x US-10-101-510-81 (1-4473)
```

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 Db 175 ATGGAGCTGGCGCTGTGGCGCTGGGGCTCTCTCGCCCTCTGGCCCCCGGAGCC 234
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 235 GCGAGCACCAAGTGTGACCGGCACAGACATGAAGTGTGGCTCCCTGCCAGTCCCGAG 294
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 295 ACCCACTGGACATGCTCGCCACCTCTACCAAGGCTGCCAGGTGGTGCAGGAACCTG 354
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
 Db 355 GAACTCACCTAGCTGCCACCAATGCCAGCTGTCTCTCCTCAGGATATCCAGGAGGTG 414
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 415 CAGGGCTAGCTGCTCATCGCTCACAAACCAAGTGAGCGAGGTCCCACTGCAGAGGTGGG 474
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 475 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGTGTACAAATGGA 534
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 535 GACCCGCTGAACAATACCACTCTGTTCACAGGGCTCCCAAGAGGCTCGGGAGCTG 594
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 Db 595 CAGCTTCGAAGCTCACAGAGATCTTGAAGAGGGGTCTTATCCAGCGGAACCCCAAG 654
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
 Db 655 CTCTGCTACAGACACAGATTTTGTGAAGGACATCTTCCAAAGAACCAACAGCTGGCT 714
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 715 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 774
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 775 GGTCCCGCTGTGTGGGAGAGAGTCTGAGGATGTACAGGCTGCAGCGCACTGTCTGT 834
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
 Db 835 GCCGTGGCTGTGCCGCTGCAAGGGCCCACTGCCCACTGCTGCTGCCATGAGCAGTGT 894
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 895 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCCCTGCCCTCCACITCAACCA 954
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 955 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACAAACACAGACAGCTTTGAG 1014
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 1015 TCCATGCCCAATCCCGAGGGCCGGTATATATTGGCGCCAGCTGTGTGACTGCTGTGCTCC 1074
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 1075 TACAACCTACTTTTACGGACGTGGGATCTGCACCCCTGCTGCCCGCTGCACAAACAA 1134
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 1135 GAGGTACACAGAGAGTGGACACAGCGGTGTGAAGTGCAGCAAGCCCTGTGCCCGA 1194
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1195 GTGTGCTATGTGGCATGAGCACTTGCAGAGAGGTGAGGCGATTACCACTGCGCAAT 1254

QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 Db 1255 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTTGGAGCTGGCATTTCTGCGGAGAGC 1314
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 Db 1315 TTTGTATGGGAGCCAGCCCTCCAAACAGTCCCGCGCTCCAGCCAGAGCAGTCCAAGTGT 1374
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 Db 1375 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCCGAGCAGCTGCCT 1434
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1435 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACAATGGCGC 1494
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1495 TACTCGCTGACCTGCAGGGCTGGGCATCAGCTGGCTGGGCTGGCTGCCTGAGGGAA 1554
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1555 CTGGCAGTGGACTGGCCCTCATCCACATACACCCACCTCTGCTTCGCACAGGGTG 1614
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1615 CCCTGGGACCACTCTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGCCA 1674
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 Db 1675 GAGGACGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGTGCCCGCCAGGGCACTGC 1734
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 Db 1735 TGGGGTCCAGGCCCCACCCAGTGTGTCAACTGCAGCCAGTTCTTCCGGGGCCAGAGTGC 1794
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 Db 1795 GTGGAGGAATCCGAGTACTGCAGGGGCTCCCAAGGAGTATGTGAATGCCAGGCACTGT 1854
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1855 TTGCGCTGCCACCCCTGAGTGTGAGCCCAAGATGGCTCAGTGCCTGTTTGGACCGGAG 1914
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 Db 1915 GCTGACCACTGTGTGGCTGTGCCCACTATAGGACCTCCCTTCTGCGTGGCCCGCTGC 1974
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 Db 1975 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2034
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 Db 2035 GCGCATGCGACGCTTGGCCCATCACTGCACCCCACTCTGTTGTGGACCTGGATGATCAAG 2094
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer- 653
 Db 2095 GGCTGCCCGCCGAGCAGCAGAGACCCCTCTGACGTCCATCATCTCTCGGGTGGTGGC 2154
 QY 653 653
 Db 2155 ATTCTGCTGCTGCTGTGGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGCAGCAG 2214
 QY 653 653
 Db 2215 AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGNAACGAGCTGTGTGGAGCCGCTG 2274
 QY 653 653
 Db 2275 ACACCTAGCGGAGCGATGCCAACCGAGCGCAGATCGCGATCTCTGAAAGACGAGAGCTG 2334
 QY 653 653

QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40		1315	TTTGTGGGAGCCAGCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1374
Db	235	GGAGACCAAGTGTGCACCGGCACAGACATGAAGCTCGCGTCCCTGCCAGTCCCGAG	294		401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60		1375	GAGACTCTCGAAGAGATCAGAGTTACCTTATACATCTCAGCATGGCCGACAGCCTGCCT	1434
Db	295	ACCCAGCTGGAGATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGGTCAGGGAACCTG	354		421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80		1435	GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGAGCAATTCTGCACATGGCGC	1494
Db	355	GAACTCACCTACTGCCCAACCAATGCCAGCCTGTCTTCTCGAGATATCAGAGGTG	414		441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100		1495	TACTCGCTCACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTCACTGAGGGAA	1554
Db	415	CAGGGCTACGTGCTCATCTCATCTCACAACTGAGGAGCAACTATGCCCTGGCCGTG	474		461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120		1555	CTGGGAGTGGACTGGCCCTCATCCACCAATAACCCACCTCTGCTTCGCACACGGTG	1614
Db	475	ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTAGACAAATG	534		481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140		1615	CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1674
Db	535	GACCCGCTGAACAATACCAACCTGTCACAGGGGCTCCCAAGAGGCTCGCGGAGCTG	594		501	GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160		1675	GAGGACGAGTGTGGCGGAGGCGCTGGCTGCCACACAGCTGTGCGCCGAGGGCACTGC	1734
Db	595	CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCA	654		521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180		1735	TGGGTTCACAGGCCACCCAGTGTGTCAACTGCACCACTTCTTCGGGGCCAGGAGTGC	1794
Db	655	CTCTGTACACAGACAGATTTTGTGAAGAGACATCTTCCACAAGAACACAGCTGGCT	714		541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200		1795	GTGGAGGAATCGCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1854
Db	715	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAAG	774		561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220		1855	TTGCGCTGCCACCTGAGTGTACGCCCCAGAAATGGTCACTGAGTACCTCTTTGGACCGGAG	1914
Db	775	GGCTCCCGTGTGGGAGAGAGTTCTGAGGATTTGTACAGAGCTGACGGGCACCTGTCTGT	834		581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240		1915	GCTGACCACTGTGTGGCCCTGTGCCCACTATAGGACCCCTCCCTCTCGTGGCCCGCTGC	1974
Db	835	GCGGTGGTGTGCCCGCTGCAAGGGGCCACCTGCCCACTGACTGTCTGCCATGAGCAGTGT	894		601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260		1975	CCACGGGTGTGAACCTGCACCTCTCTTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2034
Db	895	GCTGCGGCTGCAGCGGCCCAAGCACCTCTGACTGCTGCGCTGCCCTCCACTTCAACCA	954		621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280		2035	GGCGCATGCCAGCCTTGCCCATCAACTGCACCCACTCTGTGTGGACCTGGATGACAAG	2094
Db	955	AGTGGCATCTGTGAGCTGCACCTGCCAGGCCCTGGTCACCTACAAACACAGACACGTTGAG	1014		641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer-	653
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300		2095	GGTGTCCCGCCGAGCAGAGAGCCGCTCTGCACCTCATCATCTCTCGGTGTGTGGC	2154
Db	1015	TCCATGCCCAATCCCGAGGGCCGGTATACATTGCGGCGCAGCTGTGTGACTGCCCTGCC	1074		653	-----	653
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320		2155	ATTCTGCTGGTGTGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCCACGGCAGCAG	2214
Db	1075	TACAACCTACTTCTACGGACGTGGGATCTGCACCCCTGCTGCCCTCCGACACCAAA	1134		653	-----	653
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340		2215	AAGATCCCGAAGTACACGATGGGAGACTGTGTCAGGAAACGGAGCTGGTGGAGCCGCTG	2274
Db	1135	GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1194		653	-----	653
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360		2275	ACACCTAGCGGAGCGATGCCCAACCAAGCGCGCAGATCGCGATCCTGAAAGACAGCGAGCTG	2334
Db	1195	GTGTGCTATGGTCTGGGCATGGAGCACTTTCGAGAGGTGAGGCGAGTTACCACTGCCAAT	1254		653	-----	653
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380		2335	AGGAAGTGAAGTGTGTGGATCTGGCGCTTTTGGCACAGTCTTACAAGGGCATCTGGATC	2394
Db	1255	ATCCAGGACTTGTGCTGTCGAAGAAGATCTTTGGGAGCCTGGCATTTCTCGCGGAGAC	1314		653	-----	653
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400				

Db 2395 CCTGATGGGAGAATGTGAAATTTCCAGTGGCCATCAAAGTGTGAGGGAACACATCC 2454
 QY 653 ----- 653
 Db 2455 CCCAAAGCCAAAGAAATCTTACAGCAACATACGTGATGGCTGGTGGCTCCCA 2514
 QY 653 ----- 653
 Db 2515 TATGTCTCCCGCTTCTGGGCATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2574
 QY 653 ----- 653
 Db 2575 ATGCCCTATGGCTGCCTTTAGACCATGTCCGGGAAACCCGGGACGCTGGGTGCCAG 2634
 QY 653 ----- 653
 Db 2635 GACCTGCTGAACCTGCTGATGATGCAAGGGGATGACCTACCTGGAGGATGTGGG 2694
 QY 653 ----- 653
 Db 2695 CTCGTACACAGGGACTTGGGCGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAAA 2754
 QY 653 ----- 653
 Db 2755 ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTTGACAGACAGATACCATGCAGAT 2814
 QY 653 ----- 653
 Db 2815 GGGGCAAGGTGCCCATCAAGTGGATGGCTGGAGTCCATTCTCCGCGCGGTTCCACC 2874
 QY 653 ----- 653
 Db 2875 CACCAGAGTGTGTTGGAGTTATGGTGTGATGCTGTGGGAGCTGATGACTTTTGGGGCC 2934
 QY 653 ----- 653
 Db 2935 AAACCTTACGATGGGATCCCGCCGGAGATCCCTGACCTGCTGGAAAGGGGGAGCGG 2994
 QY 653 ----- 653
 Db 2995 CTGCCCCAGCCCCCATCTGCACCATGTGCTACATGATCATGGTCAAAATGTTGGATG 3054
 QY 653 ----- 653
 Db 3055 ATTGACTCTGAATGTGCGCCAAAGATTCCGGGAGTTGTTCTGAATTTCCCGCATGGCC 3114
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 Db 3115 AGGACCCCCAGCGCTTTGGTGCATCCAGAATGAGGACTTGGGCCCCAGCCAGTCCCTTG 3174
 QY 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684
 Db 3175 GACAGCACCCTTCTACCGCTCACTGCTGAGGACCATGACATGGGGACCTGGTGGATGCT 3234
 QY 685 GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly 704
 Db 3235 GAGGAGTATCTGTGTACCCAGCAGCGGCTTCTCTCTCCAGACCTGTCGCCGGCGCTGGG 3294
 QY 705 GlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyAspLeuThr 724
 Db 3295 GGCATGTGTCCACACAGCCAGCCGAGCTCAFTCTACAGGAGTGGCGGTGGGACCTGACA 3354
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 Db 3355 CTAGGGCTGGAGCCCTCTGAGAGAGAGGCCCCAGGTCTCCACTGGCACCCCTCCGAAGG 3414
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
 Db 3415 GCTGGCTCCGATGTATTGATGTTGACCTGGGAATGGGGCAGCAAGGGGCTGCAAAAGC 3474
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 Db 3475 CTCCCCACATATGCCCGCCCTCTACAGGGGTACAGTGGAGGACCCACAGTACCCCTG 3534

QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 Db 3535 CCCTCTGAGACTGATGGCTAGCTTGGCCCCCTGACCTGACCTGACGCCCCAGCTGAATATGTG 3594
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
 Db 3595 AACGAGCCAGATGTTCGGCCCCAGCCCTTTCGCCCCGAGAGGGGCCCTTCCCTGCTGCC 3654
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
 Db 3655 CGACCTCTGTGGTGCACCTCTGGAAGGCCCAAGACTCTCTCCACAGGGAAGATGGGCTC 3714
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
 Db 3715 GTCAAGACGCTTTTGGCTTGGGGTGCCTGGAGAACCCCGAGTACTTGACACCCAG 3774
 QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3775 GGAGGAGTGGCCCTCAGCCCCACCTCTCTGCTTTCAGCCAGCCCTTCGACAACTC 3834
 QY 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
 Db 3835 TATTACTGGGACAGGACCCACAGAGGGGGGCTCCACCCAGCACCTTCAAGGGGACA 3894
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3895 CCTACGCGAGAGAACCCAGAGTACTGGTCTGGAGCTGCCAGTG 3939

RESULT 7
 US-10-207-655-44
 ; Sequence 44, Application US/10207655
 ; Publication No. US20030118592A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ledbetter, Jeffrey A.
 ; APPLICANT: Hayden-Ledbetter, Martha S.
 ; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
 ; FILE REFERENCE: 390069.401C1
 ; CURRENT APPLICATION NUMBER: US/10/207,655
 ; NUMBER OF SEQ ID NOS: 426
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 44
 ; LENGTH: 4473
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-207-655-44

Alignment Scores:
 Pred. No.: 0 Length: 4473
 Score: 4900.00 Matches: 919
 Percent Similarity: 73.23% Conservative: 0
 Best Local Similarity: 73.23% Mismatches: 0
 Query Match: 96.49% Indels: 336
 DB: 14 Gaps: 1

SEQ6 (1-919) x US-10-207-655-44 (1-4473)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
 Db 175 ATGGAGCTGGCGGCTTGTGCGCTGGGGGTCTCTCTCGCCCTCTTGGCCCCGGAGCC 234
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 235 GCGAGCACCCAAAGTGTGCACCGCACACATGAAGCTGCGGCTCCCTTGCAGTCCCGAG 294
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 295 ACCACCTGGACATGCTCCGCCACCTCTACCGGGCTGCCAGGTGGTGCGAGGAAACCTG 354
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 355 GAATCACCCTACCTGCGCCCAATGCCAGCGCTGCTCTTCTTCCAGGATATCCAGGAGGTG 414

QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	415	CAGGGCTAGTGCTCATCGCTCACACCAAGTGGAGCAGTCCACCTGCGAGGCTGCGG	474
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	475	ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCGCTGCTAGCAATGGA	534
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyLeuArgGluLeu	140
Db	535	GACCCCTGAACAATACCACTCTGTACAGGGGCTCCCGAGGAGCTGCGGAGCTG	594
QY	141	GlnLeuArgSerLeuThrGlnIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
Db	595	CAGCTTCAGAGCTCACACAGATCTTGAAGGAGGGTCTTATCCAGCGGAAACCCAG	654
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	180
Db	655	CTCTGTACAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACACAGCTGGCT	714
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	715	CTCACACTATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG	774
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	775	GGCTCCCGTGTGGGAGAGAGTCTGAGGATGTTCAGAGCTGACGCGCACTGCTGT	834
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	835	GCGGTGGTGTGTCGCGCTGCAAGGGGCACTGCCCACTGCTGCTGCCATGAGCAGTGT	894
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	895	GCTGCCGGTGCACGGGCCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCAACCA	954
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	955	AGTGGCATCTGTGAGCTGACCTGCCAGCCCTGGTCACTACAAACACAGACAGTTTGAG	1014
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	1015	TCATGCCCAATCCCGAGGGCGGTATACATTGGCGCCAGCTGTGTGACTGCCGTGCC	1074
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1075	TACAACTACTTCTACGACGTGGATCCTGCACCCCTGCTGCCCCCTGCACAAACCAA	1134
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1135	GAGGTGACAGCAGAGGATGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1194
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1195	GTGTGTATGTGTGGGCATGGAGCACTTGCAGAGGTGAGGAGGTACCAAGTGCAT	1254
QY	361	IleGlnIlePheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1255	ATCCAGAGTTTGTGCTGCGTGCAGAAGATCTTTGGAGGCTGGCATTTCTGCGGGAGAGC	1314
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1315	TTTGATGGGAGCCAGCCCTCAACACTGCCCGCTCCAGCCAGCAGCTCCCAAGTGT	1374
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1375	GAGACTCTGGAAGAGATCAGAGTTTACCTATACATCTCAGCATGGCCGACAGCTGCT	1434
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1435	GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGAGCAATTCGCACAAATGGCGCC	1494
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1495	TACTCGCTACCTGCAAGGGCTGGGCATCAGCTGCTGGGCTGCGCTCCTCCTCAGGAA	1554
QY	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1555	CTGGGCACTGGAGCTGGCCCTCATCCACCAATAACACCCACCTCTGCTGCTGCACACGGTG	1614
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1615	CCCTGGAGCAGCTCTTTTCGGAACCCGACCAAGCTCTGCTCCACTGCCAACCCGCCA	1674
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1675	GAGGAGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGCTGCCCGCAGGAGCACTGC	1734
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1735	TGGGTTCAGGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGCGCCAGGAGTGC	1794
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1795	GTGGAGGAATCGGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1854
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1855	TTGCCGTGCCACCTGAGTGTCAAGCCCAAGATGGCTCAGTGACCTCTTTTGGACCGAG	1914
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1915	GCTGACCACTGTGGCTGTGCCACTATAAGGACCCCTCCCTTCGCTGGCGCCGCTGC	1974
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1975	CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTCGAAGTTTCCAGATGAGGAG	2034
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2035	GGCGCATGCCAGCTTGGCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAG	2094
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer-	653
Db	2095	GGTGTCCCGCGAGCAGAGAGCCCTCTGACGTCCATCATCTCTGCGGTGTGGC	2154
QY	653	-----	653
Db	2155	ATTCTGCTGCTGCTGTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGAGCAG	2214
QY	653	-----	653
Db	2215	AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGAAACGAGGCTGGTGGAGCGCTG	2274
QY	653	-----	653
Db	2275	ACACCTAGCGGAGCGATGCCCAACAGCGCCAGATCGGATCCTGAAGAGAGCGAGCTG	2334
QY	653	-----	653
Db	2335	AGGAAGGTGAAGGTGCTTGGATCTGCGCTTTTGGCACAGCTTACAAGGGCATCTGGATC	2394
QY	653	-----	653
Db	2395	CCTGATGGGAGAAATGTGAAATTCAGTGGCCATCAAAAGTGTGTAGGGGAAAAACATCC	2454
QY	653	-----	653
Db	2455	CCCAAGCCACAAGAAATCTTTAGACGAAGCATACGTGATGGCTGGTGTGGGCTCCCCA	2514
QY	653	-----	653
Db	2515	TATGTCTCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2574
QY	653	-----	653

Db 2575 ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCCGCGACGCTTGGGCTCCAG 2634
 Qy 653 ----- 653
 Db 2635 GACCTGCTGAAGTGTATGACATGTCGAAGGGGATGAGCTACCTGGAGGATGTCGG 2694
 Qy 653 ----- 653
 Db 2695 CTCGTACACAGGGACTTGGCGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAA 2754
 Qy 653 ----- 653
 Db 2755 ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATGCGAGAT 2814
 Qy 653 ----- 653
 Db 2815 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCCGCGCGGTTCCAC 2874
 Qy 653 ----- 653
 Db 2875 CACCAGAGTATGTGTGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2934
 Qy 653 ----- 653
 Db 2935 AAACCTTACGATGGATCCAGCGCCGGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2994
 Qy 653 ----- 653
 Db 2995 CTGCCCCAGCCCCCTCTGCACCATTTGATGTCTACATGATCATGTGTCAAATGTTGGATG 3054
 Qy 653 ----- 653
 Db 3055 ATTGACTCTGAATGTCTGGCCCAAGATTTCCGGAGTTGCTGTGAATTTCTCCGCGATGGCC 3114
 Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 Db 3115 AGGGACCCCGAGCGCTTTGTGGTCAATCCAGAATGAGGACTTGGCGCCAGCAGTCCCTTG 3174
 Qy 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684
 Db 3175 GACAGCACTTCTACCGCTCCTGCTGGAGAGCATGACATGGGGGACCTGGTGATGCT 3234
 Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
 Db 3235 GAGGAGTATCTGTACCCAGCAGGCGTTCTTCTCCAGACCTGCGCCGGCGCTGGG 3294
 Qy 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
 Db 3295 GGCATGCTCCACACAGGCACCGCTCATCTACCAAGAGTGGCGTGGGACCTGACA 3354
 Qy 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 Db 3355 CTAGGGCTGAGCGCTCTGAAGAGGAGGCGCCCGAGGCTCCACTGGCACCTCCGAAGG 3414
 Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeuGlnSer 764
 Db 3415 GCTGGCTCCGATGTAATGATGCTGACCTGGGAGTGGGCGACCGCAGGGGCTGCAAGC 3474
 Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 Db 3475 CTCCCCACATGACCCCGCCCTTACAGCGGTACAGTGGAGACCCACAGTACCCCTG 3534
 Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 Db 3535 CCCTCTGAGACTGATGGCTACGTTGGCCCTTACCTGTCAGCGCCCGAGCGCTGAATATG 3594
 Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
 Db 3595 AACCAGCCAGATGTTGGCCCCAGCCCCCTTCCGCCGAGAGGCGCTCTGCTGCTGCC 3654
 Qy 825 ArgProAlaGlyAlaThrLeuGluArgProProLysThrLeuSerProGlyLysAsnGlyVal 844

Db 3655 CGACCTGCTGGTGCCACTCTGGAAGGCCCAAGACTCTTCCCGAGGAAGAATGGGTC 3714
 Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
 Db 3715 GTCAAGACGCTTTTGGCTTGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCAG 3774
 Qy 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3775 GGAGGAGCTCCCTCAGCCCACTCTCTGCTGCTTCCAGCCAGCGCTTCGACACCTC 3834
 Qy 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
 Db 3835 TATTACTGGACGAGGACCCAGAGCGGGGCTCCACCCAGCACCTTCAAAGGACA 3894
 Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3895 CCTACGGCAGAGAACCAGAGTACCTGGGTCTGGACGTGCCAGTG 3939

RESULT 8
 US-10-207-498-5
 ; Sequence 5, Application US/10207498
 ; Publication No. US20030143568A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Elizabeth Singer
 ; APPLICANT: Ralf Landgraf
 ; APPLICANT: Dennis J. Slamon
 ; APPLICANT: David Eisenberg
 ; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
 ; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREGULIN AND HER3
 ; FILE REFERENCE: 30448.103-US-01
 ; CURRENT APPLICATION NUMBER: US/10/207,498
 ; PRIOR FILING DATE: 2002-07-29
 ; PRIOR APPLICATION NUMBER: 60/308,431
 ; PRIOR FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 3765
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(3765)
 ; US-10-207-498-5

Alignment Scores:
 Pred. No.: 0 Length: 3765
 Score: 4892.00 Matches: 918
 Percent Similarity: 73.15% Conservative: 0
 Best Local Similarity: 73.15% Mismatches: 1
 Query Match: 96.34% Indels: 336
 DB: 12 Caps: 1

SEQ6 (1-919) x US-10-207-498-5 (1-3765)

Qy 1 MetGluLeuAlaAlaLeuCysArgTyrGlyLeuLeuAlaLeuLeuProGlyAla 20
 Db 1 ATGAGCTGGCGGCTTTGGCCCTGGGGCTCTCTCTCGCCCTCTTGGCCCCCGGAGCC 60
 Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 61 GCGAGCACCAAGTGTGCACCGCAGACATCAAGCTGGGCTCCCTGCCAGTCCCGAG 120
 Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 121 ACCACCTGGACATGCTCCGCCACCTCTACCAAGGCTGGCGAGTGGTGCAGGAAACCTG 180
 Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspGlnGlnVal 80
 Db 181 GAACCTCACCTACCTGCGCCCAATGCGAGCTCTCTCTCTGAGGATATCCAGGAGGTG 240
 Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100

Db 241 CAGGGCTACGTCTCATCGCTCACAAACCAGGTGAGCGAGGTCCCTCCACTGCAGAGGCTGCGG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTACAAATGGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 361 GACCCGCTGAACAATACCAACCCTGTTCACAGGGGCTCCCCAGGAGGCTGCGGAGCTG 420
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 Db 421 CAGCTCGAAGCTCACAGAGATCTTGAAGAGGGGTCTTATCATCAGCGAACCCTCCAG 480
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 Db 481 CTCCTGTACAGACACGATTTTGTGAAGGACATCTTCCACGAACAACACAGCTGGCT 540
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCystLys 200
 Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600
 QY 201 GlySerArgCysTrpGlyIleSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 601 GGTCTCCGCTGTGGGAGAGATCTGTAGGATGTCAGAGCCTGACGCGCACTGTCTGT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 Db 661 GCCGTGTGTGCCCCGTGCAAGGGCCACTGCCCACTGACTGTGCCATGAGCAGTGT 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 721 GCTGCCGCTGACGGGCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCAACCCAC 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 781 AGTGGCATCTGTGAGCTGCCTGCCAGCCCTGGTCACTACACACAGACAGTTTGAG 840
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 841 TCCATGCCCAATCCCGAGGGCCGTTATACATTCGGCGCCAGCTGTGTGACTGCCGTGCC 900
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 901 TACAACACTTCTTACGGACCTGGGATCTCTCACCCCTGCTGCCCCCTGCACAAACCAA 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 961 GAGGTGACAGCAGAGGATGGAACACAGCCGTTGTGAAGTGCAGCAAGCCCTGTGCCCGA 1020
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1021 GTGTGCTATGCTGGGCATGGAGCACTTGCAGAGGTGAGGCGACTTACCAAGTGCAT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 Db 1081 ATCCAGGAGTTTGTGCTGGCTGCAAGAAGATCTTTGGAGCCTGGCAATTTCTGCGGAGAGC 1140
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400
 Db 1141 TTTGATGGGAGCCAGCCCTCCAACTGCTCCCGCTCCAGCCAGCAGCTCAAGTGT 1200
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 Db 1201 GAGACTCTGGAAGAGATCACAGTTTACCTATATACATCTCAGCATGGCCGAGAGCTGCCT 1260
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1261 GACCTCAGGCTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATCTGCACAATGGCGCC 1320
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1321 TACTCCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCACTGAGGGAA 1380

QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1381 CTGGGCAGTGGACTGGCCCTCATCCACCATAACCCACCTCTGCTTCTGTGCACAGGTG 1440
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1441 CCCTGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACTGCCAACCGGCA 1500
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 Db 1501 GAGGACGAGTGTGTGGCGAGGCTGGCTGCCACCACTGTGCGCGGAGGACACTGC 1560
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 Db 1561 TGGGGTCCAGGGCCCAACCCAGTGTCTCACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 Db 1621 GTGGAGGAATCCGAGTACTGCAGGGGTCTCCAGGGAGTATGTGAATGCCAGGCACTGT 1680
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1681 TTGCCGTGCCACCCCTGAGTGTGAGCCCAAGATGGCTCAGTGCCTGTTTGGACCGGAG 1740
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 Db 1741 GCTGACCACTGTGGCTGTGCCCTGTGCCCACTATAGGACCCTCCCTTCTGCGTGCCTGC 1800
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 Db 1801 CCCAGCGGTGTGAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 Db 1861 GCGCATGCCACCTTGGCCCATCAACTGCACCCCACTCTCTGTGTGGACCTGGATGACAAG 1920
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
 Db 1921 GGCTGCCCGCGGAGCAGAGAGCCGCCCTCTGACGTCCATCTCTCGCGTGGTGTGGC 1980
 QY 653 ----- 653
 Db 1981 ATTCTGCTGCTCGTGGTCTTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGAG 2040
 QY 653 ----- 653
 Db 2041 AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAACGAGCTGGTGGAGCGCGTG 2100
 QY 653 ----- 653
 Db 2101 ACACCTAGCGGAGCGATGCCCAACGAGCGCAGATGCGGATCCTGAAAGAGACGAGGAGTG 2160
 QY 653 ----- 653
 Db 2161 AGGAAGGTGAAGGTGCTTGGGATCTGGGCGCTTTTGGACAGTCTTACAAAGGCATCTGGATC 2220
 QY 653 ----- 653
 Db 2221 CCTGATGGGGAATGTGAAAATTTCAGTGGCCATCAAAAGTTTGAGGGAACACATCC 2280
 QY 653 ----- 653
 Db 2281 CCCAAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGGGCTCCCCA 2340
 QY 653 ----- 653
 Db 2341 TATGTCTCCCGCTTCTGGGCACTCTGCCTGACATCCACGGTGCAGCTGGTGTGACACAGT 2400
 QY 653 ----- 653
 Db 2401 ATGCCCTATGGCTGCTCTTAGACCATGTCCCGGAAACCGCGGAGCCCTGGGCTCCCAG 2460

Qy 653 ----- 653
 Db 2461 GACCTGTGAACCTGGTGTATGACAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGG 2520
 Qy 653 ----- 653
 Db 2521 CTCGTACACAGGGACTTGGCGGCTCGGAACGTCGTGTCAAGAGTCCCAACCACTGTCAAA 2580
 Qy 653 ----- 653
 Db 2581 ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATGCAGAT 2640
 Qy 653 ----- 653
 Db 2641 GGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGCTTCACC 2700
 Qy 653 ----- 653
 Db 2701 CACCAGAGTATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2760
 Qy 653 ----- 653
 Db 2761 AAACCTTACGATGGATCCCGAGCCCGGAGATCCCTGACCTGTGGAAGGGGGAGCGG 2820
 Qy 653 ----- 653
 Db 2821 CTGCCCGAGCCCCCATCTGCACCATTTGATGCTTACATGATCATGTGTCATGTCGATG 2880
 Qy 653 ----- 653
 Db 2881 ATTGACTCTGAATCTCGGCCAAGATTCCGGAGTGTGTGTGAATTTCTCCCGCATGGCC 2940
 Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 Db 2941 AGGACCCCCAGCCCTTTGGTTCATCCAGATGAGGACTTGGCCCGCAGCAGTCCCTTG 3000
 Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 Db 3001 GACAGCACCTTCTACCCCTCCTGCTGAGGAGCATGACATGGGGGACCTGGTGGATGCT 3060
 Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
 Db 3061 GAGGAGTATCTGGTACCCAGCAGGAGGCTTCTGTCTCAGACCCCTGCCCGGGCGCTGG 3120
 Qy 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
 Db 3121 GGCATGCTCCACACAGCAGCCGAGCTCATCTACAGAGAGTGGCGGTGGGGACCTGACA 3180
 Qy 725 LeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 Db 3181 CTAGGGCTGGAGCCCTCTGAAGAGAGGAGGCCCGCCAGGTCTCCACTGGCACCCCTCCGAAGG 3240
 Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
 Db 3241 GCTGGCTCCGATGATTGATGGTGAAGTGGGAAATGGGGCAGCAGAGGGGCTGCCAAGC 3300
 Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 Db 3301 CTCCCCACACATGACCCCGCCCTCTACAGCGGTACAGTACAGACCCCGCAGTACCCCTG 3360
 Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 Db 3361 CCCTCTGAGACTGATGGCTACGTTGCCCTGACCTGACGCGCCCGCCAGCCCTGAATATGTG 3420
 Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
 Db 3421 AACAGCCAGATGTTGGCCCGCCAGCCCTTCCCGCGAGAGGGCCCTCTGCTGCTGCC 3480
 Qy 825 ArgProAlaGlyAlaThrLeuGluArgProGlyThrLeuSerProGlyLysAsnGlyVal 844
 Db 3481 CGACCTGCTGGTCCACTCTGGAAAGGGCCAGACTCTCTCCCGCAGGGAAGATGGGGTC 3540
 Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864

Db 3541 GTCAAAACAGCTTTTGGCTTTGGGGTGGCCGTGGAGAACCCCGAGTACTTTGACACCCAG 3600
 Qy 865 GlyGlyAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3601 GGAGGACTGCCCTCAGCCCCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3660
 Qy 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
 Db 3661 TATTACTGGGACAGGACCCACAGAGGGGGGCTCCACCCAGCACCTTCAAGAGGACA 3720
 Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3721 CCTACGCGACAGAACCCAGAGTACTGGGTCTGGAGCGTGCCAGTG 3765

RESULT 9
 US-09-811-123-8
 ; Sequence 8, Application US/09811123
 ; Patent No. US20020001587A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sharon Erickson
 ; APPLICANT: Ralph Schwall
 ; APPLICANT: Mark Slikowski
 ; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
 ; FILE REFERENCE: ANTIBODY-MAYTANSINOID CONJUGATES
 ; FILE REFERENCE: GENENT.073A2
 ; CURRENT FILING DATE: 2001-03-16
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR FILING DATE: 2000-10-05
 ; PRIOR FILING DATE: 09/602,530
 ; PRIOR FILING DATE: 2000-06-23
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 3768
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-811-123-8

Alignment Scores:
 Pred. No.: 0 Length: 3768
 Score: 4892.00 Matches: 918
 Percent Similarity: 73.15% Conservative: 0
 Best Local Similarity: 73.15% Mismatches: 1
 Query Match: 96.34% Indels: 336
 DB: 9 Gaps: 1

SEQ6 (1-919) x US-09-811-123-8 (1-3768)
 Qy 1 MetGluLeuAlaAlaLeuCysArgTyrGlyLeuLeuLeuAlaLeuLeuProGlyVala 20
 Db 1 ATGAGCTGGCGGCTTGTCCGCTGGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
 Qy 21 AlaserThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 61 GCGAGCACCACCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 120
 Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 121 ACCACCTGGACATGTCTCCGCCACCTCTACAGGGCTGCCAGGTGTGTGTGTGTGTGTGTGT 180
 Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 181 GAACTCACCCTACCTGCTCCCAATGCCAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
 Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 241 CAGGGCTACGTGCTCATCGCTCACACCAAGTAGGAGGAGGTCCCACTGCAGAGGCTCGCG 300
 Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 301 ATTTGCGAGGCCACCCAGCTCTTTTGGAGCAACATATGCCCTGGCGCTGTAGACAATGGA 360


```
Db 2521 CTCGTACACAGGGACTTGGCGCTCGGAACGTCGTGTTCAAGAGATCCCAACCACTGTCAAA 2580
Qy 653 ----- 653
Db 2581 ATTACAGACTTCGGGTGGCTCGGCTGCTGACATTTGACGACAGAGATACCATGCAGAT 2640
Qy 653 ----- 653
Db 2641 GGGGCAAGGTGCCCAATCAAGTGGATGGCTGGAGTCCCAATCTCCGCGCGCGTTCACC 2700
Qy 653 ----- 653
Db 2701 CACCAGAGTGATGTGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2760
Qy 653 ----- 653
Db 2761 AAACCTTACGATGGATGCCAGCCGGGAGATCCCTGACCTGTGGAAAGGGGAGGG 2820
Qy 653 ----- 653
Db 2821 CTGCCCCAGCCCCCATCTGCACCATTTGATGCTTACATGATCATGTTGCAATGTTGGATG 2880
Qy 653 ----- 653
Db 2881 ATTGACTCTGAATCTCGGCCAAGATTCGGGAGTTGGTCTCTGAATTCTCCCGCATGGCC 2940
Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 2941 AGGACCCCCAGCCCTTGTGGTTCATCAGAAATGAGGACTTGGGCCCGCAGCTCCCTTG 3000
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCTTCTACCGCTCACTGCTGGAGACATGACATGGGGACCTGGTGGATGCT 3060
Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGGTACCCCAAGCAGGAGGCTTCTTCTCCAGACCCCTGCGCGGGCGCTGG 3120
Qy 705 GlyMetValHisArgHisArgSerSerSerThrArgSerGlyGlyGlyAspLeuThr 724
Db 3121 GGCAATGTCACACAGCCAGCCAGCTCATCTACACAGGAGTGGCGGTGGGGACCTGACA 3180
Qy 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3181 CTAGGGCTGGAGCCCTCTGAAGAGAGGAGGCCCCAGGCTCTCCACTGGCACCCCTCGAAGG 3240
Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3241 GCTGGCTCCGATGTATTTGATGGTGGCTGGGAATGGGGCAGCCAAAGGGGCTGCAAAAGC 3300
Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3301 CTCCCCACATGACCCCGCCCTCTACAGGGTACAGTACAGGACCCACAGTACCCCTG 3360
Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3361 CCCTCTGAGACTGATGGCTACGTTGCCCTGACCTGACCTGACGCCCCAGCCCTGAATATGTG 3420
Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3421 AACAGCCAGATGTTGGCCCCCAGCCCTTCGCCCCGAGAGGGCCCTCTGCTGTGCTGCC 3480
Qy 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3481 CGACCTGCTGGTCCACTCTGGAAAGGGCCCAAGACTCTCTCCCGAGGAGAGATGGGGTC 3540
Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3541 GTCAAGACGTTTTTGGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCGAG 3600
Qy 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3601 GAGGAGTGGCTTCCAGCCCCACCTCTCTCTGCTTACGCCAGCGCTTCGACCAACCTC 3660
Qy 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3661 TATTACTGGGACCAAGGACCCACCAGAGCGGGGGCTCCACCCAGCACCTTCAAAGGGACA 3720
Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3721 CCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGCAGCTGCCAGTG 3765

RESULT 10
US-09-811-115-2
: Sequence 2, Application US/09811115
: Patent No. US20020035736A1
: GENERAL INFORMATION:
: APPLICANT: Erickson, Sharon
: APPLICANT: Schwall, Ralph
: APPLICANT: King, Kathleen
: TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
: FILE REFERENCE: GENENT.034A
: CURRENT APPLICATION NUMBER: US/09/811,115
: CURRENT FILING DATE: 2001-03-16
: PRIOR APPLICATION NUMBER: 60/189,844
: PRIOR FILING DATE: 2000-03-16
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 3768
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred No.: 0 Length: 3768
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 9 Gaps: 1

SEQ6 (1-919) x US-09-811-115-2 (1-3768)
Qy 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGAGCTGGCGGCTTGTGCCGTGGGGCTCTCTCGCTCTTGGCCCCCGAGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCAAGTGTGCACCGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 120
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCCACTGGACATGTCGCCACCTCTACCAAGGCTGCCAGGTGGTGGAGGGAACCTG 180
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAACCTACCTACCTGCGCCCAATGCCAGCGCTGCTCTCTGTCAGGATATCCAGAGGTG 240
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGTGCTCATCGCTCACAAAGTACAGGAGGCTCCACATGCCAGAGGCTCGG 300
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACATGGA 360
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGCTGAACAATACCACTCCCTGTACAGGGGCTCCCGCAGGAGGCTCGCGGAGCTG 420
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTCGAAGCCTCAGAGATCTTTGAAAGAGGGGTCTTTGATCCAGCGGAACCCCGAG 480
```

QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	180
Db	481	CTCTGCTACCAAGACACAGATTTTGTGAAGGACATCTTCCACAAGAACAACAGCTGGCT	540
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	541	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG	600
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	601	GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTCAGAGCCTGACGGCCTGCTGT	660
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys	240
Db	661	GCCGCTGGCTGTGCCGCTGCAAGGGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTG	720
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	721	GCTGCCGCTGCACGGGCCCCAAGCCTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTG	780
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATCTGTAGCTGCACCTGCCAGCCCTGGTCACTACACACAGACAGCTTTGAG	840
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCCGAGGGCCGTATACATTCGGCCAGCTGTGTGACTGCCCTGCC	900
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACTACTTTTACGGACCTGGGATCCTGCACCCCTGCTGCTGCTGCTGCTGCTG	960
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACAGCAGAGATGGACACAGCGGTGTGAGAGTGAGCAAGCCCTGTGCCCGA	1020
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGTCTGGGCATGGACACTTGCAGAGGTGAGGCGAGTTACCAGTGCCAT	1080
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGAGATTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGC	1140
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400
Db	1141	TTTGATGGGGACCCAGCCTCCAACTGCCCCGCTCCAGCCAGAGAGCTCCAAGTGT	1200
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGCGCGACAGCTGCCT	1260
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGAGCAATTCGCACAATGGCGC	1320
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTGCTGCTGCTG	1380
QY	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGGCACTGGAGCTGGCCCTCATCCACCATAACACCCACCTCTGCTTGTGTGCACG	540
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCACTCTTTCGGAACCCCGCACCAAGCTGCTGCACACTGCCAACCGGCA	1500
QY	501	GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGAGAGTGTGGCGAGGGCCTGGCTGCCACCACTGCTGCTGCTGCTGCTGCTGCTG	1560

QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCAGTTCTTTCGGGGCCAGAGTGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATGCGAGTACTGCAGGGGTCCCCAGGAGATGTGAATGCCAGGCACGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTCGCCACCTGAGTGTGCCCCCAGAAATGCTCAGTCACTGTTTGGACCGGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCAGTGTGGCCTGTGCCACATATAAGGACCCCTCCCTTCTGCTGCGCCGCTG	1800
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCACGCGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Db	1861	GGCGCATGCCAGCCTGCCCCATCACTGCACCCACTCCTGTGTGGACCTGGATGACAAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCGAGCAGAGAGCCACCTCTGACGTCCATCGTCTCTGCGTGGTGGC	1980
QY	653	-----	653
Db	1981	ATTCTGCTGCTGCTGCTTGGGGTGGTCTTTTGGGATCCTCATCAAGGACGGCAGCAG	2040
QY	653	-----	653
Db	2041	AAGATCCGGAAGTACAGATGCGGAGACTGTGTCAGGAAACGAGCTGTGGAGCGCGTG	2100
QY	653	-----	653
Db	2101	ACACCTAGCGGAGCGATGCCAACGAGCGCAGATGCGGATCCTGAAAGAGACGAGCTG	2160
QY	653	-----	653
Db	2161	AGGAAGTGAAGTCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220
QY	653	-----	653
Db	2221	CCTGATGGGGAATGTGAAATTTCCAGTGGCCATCAAAGTGTGAGGGAACACATCC	2280
QY	653	-----	653
Db	2281	CCCAAAGCCAACAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGGGCTCCCCA	2340
QY	653	-----	653
Db	2341	TATGTCCTCCGCTTCTGGGCATCTGCCCTGACATCCACGGTGCAGCTGTGTGACAGCTT	2400
QY	653	-----	653
Db	2401	ATGCCCTATGGCTGCCCTCTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCCAG	2460
QY	653	-----	653
Db	2461	GACCTGCTGAAGTGGTGTATGCAGATTTGCCAAGGGGATGAGCTACTGGAGGATGTGCG	2520
QY	653	-----	653
Db	2521	CTCGTACACAGGAGTGTGGCGCTCGGAACGTGCTGGTCAAGAGTCCCCAACCATGTCAA	2580
QY	653	-----	653
Db	2581	ATTACAGACTTCGGGCTGGCTGGGCTGCTGGACATTGACGACACAGATACCATGCAGAT	2640
QY	653	-----	653

Dd 3721 CCTAGCGCAGAGACCCAGACTACCTGGGTCTGGACGTCGCCAGTG 3765

RESULT 11
US-09-877-177-11
; Sequence 11, Application US/09877177
; Publication No. US20020192652A1
; GENERAL INFORMATION:
; APPLICANT: Peter V. Danenberg et al.
; TITLE OF INVENTION: Method of determining Epidermal Growth
; FILE OF INVENTION: Factor Receptor and HER-2 New Gene Expression
; TITLE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177-11

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 10 Gaps: 1

SEQ6 (1-919) x US-09-877-177-11 (1-4530)

Qy	1	MetsLeuAlaLeuLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGluAla	20
Dd	151	ATGCAGCTGGGGCGCTTGTGCGCTGGGGGTCCTCCTCGCCCTCTTCCCCCGGAGCC	210
Qy	21	AlaserThrGlnValCysThrGlyThrAspMetIysLeuArgLeuProAlaSerProGlu	40
Dd	211	GCGAGCACCCCAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCCTGCCAGTCCCCGAG	270
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
Dd	271	ACCACCTGGACATGCTCCGCCACCTCTACACAGGGCTGCCAGGTGGTCAGAGGAACCTG	330
Qy	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal	80
Dd	331	GAATCATCTACCTACCTGCCACCACAATGCCAGCTGTCTCTTCTGCAGGATATCCAGGAGGTG	390
Qy	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Dd	391	CAGGGCTACGTGCTCATCGCTCAACAACCAAGTAGAGCAGGTGCCACTGCAGAGGTGGG	450
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Dd	451	ATTGTGGAGGCACCCAGCTCTTTGAGGACAACATATGCCCTGGCCGCTGCTAGACAATGGA	510
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Dd	511	GACCCGCTGAACAATACACCCCTGTCCACAGGGGCTCCCCAGAGGAGGCTCGGGGAGCTG	570
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Dd	571	CAGCTTCGAAGCCTCACAGAGATCTTAAAGAGGGGTCTTGATCCAGCGGAACCCCCAG	630
Qy	161	LeuCystyrGlnAspThrIleLeuTrpIysAspIlePheHisLysAsnAsnGlnLeuAla	180
Dd	631	CTCTGCTACCAGGACACGATTTTGTGAAGAGACATCTTCCAAGAACAACACCGCTGGCT	690
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Dd	691	CTCACACTGATAGACCAACACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAAG	750
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220

Db	751	GGCTCCCGCTGCTGGGAGAGAGTCTCTGAGGATTGTGAGAGCCTGACGGCACACTGTCTGT	810	QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240	Db	1891	GCTGACCACTGTGTGCCCTGTGCCACTATAAGGACCCCTCCCTTCTGCGTGGCCGCTGC	1950
Db	811	GCGGTGGCTGTGCGCGCTGCRAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT	870	QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	Db	1951	CCGAGGGTGTAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2010
Db	871	GCTGCCGGTGCACGGGCCCAAGCACTCTGACTGCTGCCCTGCCTCCACTTCAACCCAC	930	QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	Db	2011	GGCGCATGCCAGCTTGGCCCCATCAACTGCACCCACTCTCTGTGTGACCTGGATGACAA	2070
Db	931	AGTGGCATGTGAGTGCACCTGCCAGCCCTGGTCACCTACAACACACACACGCTTTGAG	990	QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	Db	2071	GGCTGCCCGCCGAGCAGAGCCAGCCCTCTCGACGTCCATCTCTGCGGTGGTTGGC	2130
Db	991	TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCC	1050	QY	653	-----	653
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	Db	2131	ATTCTGCTGCTGCTGCTTGGGGTGGTCTTTTGGGATCCTCATCAAGGACGCGCAGCAG	2190
Db	1051	TACAACCTAGCTTTCTACGACGTGGGATCTGCACCCCTGCTCTGCCCTGCACAAACCAA	1110	QY	653	-----	653
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	Db	2191	AAGATCCGAAAGTACAGATGGGAGACTGCTGCAGAAACGAGAGCTGTGGAGCCGCTG	2250
Db	1111	GAGGTACACACAGAGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1170	QY	653	-----	653
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	Db	2251	ACACCTAGCGGAGCATGCCAACACAGCGCAGATCGGGATCCTGAAAGAGACGGAGCTG	2310
Db	1171	GTGTGCTATGGTCTGGGCATGAGCACTTGCAGAGGTGAGGCACTTTACAGTGCCTAAT	1230	QY	653	-----	653
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	Db	2311	AGGAAGGTGAAGTGTGATCTGGCTTTTGGCACAGTCTACAAAGGCGCATCTGGATC	2370
Db	1231	ATCCAGAGTGTGCTGGCTGCAAGAAGATCTTTGGGAGCTGGCATTTCTGCCGGAGAGC	1290	QY	653	-----	653
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400	Db	2371	CCTGATGGGAGAAATGTGAAATTCAGTGGCCATCAAAGTTGTGAGGGAACACATCC	2430
Db	1291	TTTGATGGGACCCAGCCCTCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1350	QY	653	-----	653
QY	401	GluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420	Db	2431	CCAAAGCCAACAAGAAATCTTAGACGAAGCATACGTGATGCTGTTGGGCTCCCCA	2490
Db	1351	GAGACTCTGGAGAGATCAGAGTTACCTATACATCTCAGCATGGCGCGACAGCTGCCT	1410	QY	653	-----	653
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	Db	2491	TATGTCCTCCGCTTCTGGGCACTCTGCCCTGCATCCACGTCGAGCTGTGTGACACAGCTT	2550
Db	1411	GACCTCAGCGTCTTCCAGAACCTTCAAGTAATCCGGGGAGCAATTCGCACAATGGCGCC	1470	QY	653	-----	653
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460	Db	2551	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCCAG	2610
Db	1471	TACTCGCTGACCTGCACGGCTGGGCATCAGCTGGCTGGGGCTGCCTCACTGAGGGAA	1530	QY	653	-----	653
QY	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480	Db	2611	GACCTGCTCAACTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2670
Db	1531	CTGGCAGTGGACTGGCCCTATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTG	1590	QY	653	-----	653
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	Db	2671	CTCGTACACAGGACTTGGCCGTCGGAACGTGCTGGTCAAGAGTCCCCAACCATGTCAAA	2730
Db	1591	CCCTGGGACAGCTCTTGGGAACCGCCACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1650	QY	653	-----	653
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	Db	2731	ATTACAGACTTCGGGCTGGCTCGGCTGGACATTGACGACACAGTACCATTGCAGAT	2790
Db	1651	GAGCAGAGTGTGGGCGAGGCGCTGGCCCTGCCACAGCTGTGCCCGCCGAGGCACTGC	1710	QY	653	-----	653
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540	Db	2791	GGGGCAAGGTGCCCCATCAAGTGGATGGCGCTGGAGTCCATCTCCGCCGCGGTTTACC	2850
Db	1711	TGGGTCCAGGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1770	QY	653	-----	653
QY	541	valGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560	Db	2851	CACCAAGATGATGTGTGGAGTTATGCTGTACTGTGTGGAGCTGATGACTTTTGGGGCC	2910
Db	1771	GTGGAGGAATTCGGAGTACTGCAAGGGCTCCCGAGGGAGTATGTGAATGCCAGGCATGT	1830	QY	653	-----	653
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580	Db	2911	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGTGAAAGGGGAGCGG	2970
Db	1831	TTGCCGTGCCACCTCAGTGTGACGCCCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG	1890				

```
Qy 653 ----- 653
Db 2971 CTGCCCGAGCCGCCATCTGCACCATTTGATGCTACATGATGCTCAAATGTTGGATG 3030
Qy 653 ----- 653
Db 3031 ATTGACTCTGAATCTCGGCCAAGATTCGGGAGTTGCTGTGAATTTCTCCCGCATGGCC 3090
Qy 654 -----
Db 3091 AGGACCCCGAGCCGTTTGGTTCATCCAGAAATGAGACTTGGCCCGACGCTCCCTTG 3150
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACCTTCTACCGCTCCTGAGGACGATGACATGGGGACCTGGTGGATGCT 3210
Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProaspProAlaProGlyAlaGly 704
Db 3211 GAGGAGTATCTGGTACCCAGCAGGCGCTTCTTGTCCAGACCTGCCCCGGGCGCTGGG 3270
Qy 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
Db 3271 GGCATGTCTCACACGACGACCGAGCTCATCTACAGGAGTGGCGGTGGGACCTGACA 3330
Qy 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3331 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCGCCAGGCTCTCCACTGGCACCTCCGAAGG 3390
Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3391 GCTGGCTCCGATGTATTGATGTGACCTGGGATGGGAGCGAGCCAAAGGCGTCAAGC 3450
Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3451 CTCCCAACATGACCCCGCCCTCTACAGCGGTACAGTGGAGGCCCGCCAGTACCCCTG 3510
Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3511 CCCTCTGAGACTGATGCTAGCTGCCCCCTGACCTGCAGCCCGACGCTGAATATGTG 3570
Qy 805 AsnGlnProaspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3571 AACCAGCAGATGTTGGGCCCGCCAGCCCTTCCGCCGAGAGGGCCCTCTGCCTGTGCTC 3630
Qy 825 ArgProAlaGlyAlaThrLeuGluArgProIysThrLeuSerProGlyLysAsnGlyVal 844
Db 3631 CGACCTCTGGTGCATCTCGAAAGGCGCCAAAGACTCTCTCCCGAGGAGGAGGAGTGGGTC 3690
Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3691 GTCAAGACGTTTGTGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCCGAG 3750
Qy 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3751 GGAGGAGCTGCCCTCAGCCCGCCAGCCCTCTCTGCTTCCAGCCAGCGCTTCGACAACTC 3810
Qy 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3811 TATTACTGGACGAGGAGGCCAGCAGCGGGGGGCTCCACCCAGCAGCTTCAAGAGGACA 3870
Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3871 CCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG 3915
```

RESULT 12

```
US-10-007-926A-119
; Sequence 119, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
```

```
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1346-R-00
; CURRENT APPLICATION NUMBER: US/10/007.926A
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: v-erb-b2 avian erythroblastic leukemia viral
; OTHER INFORMATION: oncogene homolog 2 (neuro/glioblastoma derived
; OTHER INFORMATION: oncogene homolog) (ERBB2) gene.
US-10-007-926A-119
```

Alignment Scores:

```
Pred. No.: 0 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 12 Gaps: 1
```

SEQ6 (1-919) x US-10-007-926A-119 (1-4530)

```
Qy 1 MetGluLeuAlaAlaLeuCysArgTyrGlyLeuLeuLeuAlaLeuProGlyAla 20
Db 151 ATGAGCTGGCGGCTTGTGCCCTGGGGCTCTCTCGCCCTCTTCCGCCCGGAGCC 210
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGln 40
Db 211 GCGAGCACCAAGTGTCCACCGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 270
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGCTCCGCCACCTCTACGAGGCTGCCAGGTGGTGCAGGGAACCTG 330
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspLeuGlnVal 80
Db 331 GAACCTACCTACCTGCCCGCCCAATGCCAGCTGTCTCTCTCGAGGATATCCAGAGGTG 390
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACGTGCTCATCGCTCACAACTGAGGAGGCTCCCACTGCAGAGCTCGCG 450
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGGAGGACCCAGCTCTTTGAGGACAACCTATGCCCTGGCGCTGCTAGACAATGA 510
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAAATACCACTCCCTGTCCAGGGGCTCCCGAGGAGGCTCGCGGAGCTG 570
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTCGAAGCCTCACAGAGATCTTGAAGGAGGCTCTTATCCAGCGAGGAGGAGGAG 630
Qy 161 LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGCTACCCAGGACACGATTTTGTGAAGGACATCTTCCACAAGAAACACAGCTGGCT 690
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCCGCTCTCGGGGCTTCCACCCCTGTCTTCCGATGTGTAAG 750
Qy 201 GlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 201 GlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
```

Db	751	GGTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTTGTCAGAGCCTGACGGCAGCTGTCGT	810	QY	581	AlaAspGlnCysValalaCysAlaHisTyrIysAspProPheCysValalaArgCys	600
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240	Db	1891	GCTGACCAGTGTGGCCTGTGCCACTATAAGGACCCCTCTCTGCTGCGCCGCTGC	1950
Db	811	GCCGCTGGCTGTGCCGCTGCAAGGGCCACTGCCACTGACTGCTGCCATGACAGTGT	870	QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	Db	1951	CCACGGGTGTGAACCTGACCTCTCATATGCCATCTGGAAGTTTCCAGATGAGAG	2010
Db	871	GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCTGGCCTGCCCTCCACTCAACAC	930	QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	Db	2011	GCGCATGCCAGCCTTGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	2070
Db	931	ATGGCATCTGTAGCTGCACCTGCCAGCCCTGGTCACCTACACACACACACGTTTGA	990	QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	Db	2071	GGCTCCCGCGCAGCAGAGAGCCAGCCCTCTGACGTCCATCTCTCTGCGGTGTGGC	2130
Db	991	TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCCTCC	1050	QY	653		653
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	Db	2131	ATTCTGCTGCTGCTGCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGAG	2190
Db	1051	TACAACTACCTTTCTACGGACCTGGGATCTCTGCACCTCGTCTGCCCTCGCACACCAA	1110	QY	653		653
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	Db	2191	AGATCCGGAAGTACACAGTACCGGAGACTGCTGCAGGAACCGAGCTGCTGGAGCCGCTG	2250
Db	1111	GAGGTGACACAGAGAGTGAACACAGCGGTGTGAGAGTGCAGAACGCCCTGTGCCCGA	1170	QY	653		653
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	Db	2251	ACACCTAGCGGAGCGATGCCAACAGCGCGCAGATGCGGATCCTGAAAGAGACGGAGCTG	2310
Db	1171	GTGTGCTATGGTCTGGGCATGGACACTTGCAGAGGTGAGGCGAGTTACCACTGCCAAT	1230	QY	653		653
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	Db	2311	AGGAAGTGAAGTCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2370
Db	1231	ATCCAGGAGTTTGTGGCTGCAAGAGATCTTTGGAGCCTGGCATTTCTGCCGGAGAGC	1290	QY	653		653
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGlnLeuGlnValPhe	400	Db	2371	CCTGATGGGGAATGTGAAATTCAGTGGCCATCAAAAGTGTGAGGGAACACACATCC	2430
Db	1291	TTTGATGGGACCCAGCCTCAACACTGCCCGCTCCAGCCAGACAGCTCCAAAGTGT	1350	QY	653		653
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420	Db	2431	CCCAAAGCAACAAGAAATCTTAGACGAAGCATACGTGAGTGGCTGTGGGCTCCCCA	2490
Db	1351	GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGCGCGACAGCCTGCCT	1410	QY	653		653
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	Db	2491	TATGTCCTCCGCTTCTGGGCATCTGCTGACATCCACGGTGACGCTGCTGACACAGCTT	2550
Db	1411	GACCTCAGCGTCTTCAGAACCTGCAAGTAATCCGGGACGAAATCTGCACAAATGGCGC	1470	QY	653		653
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyArgSerLeuArgGlu	460	Db	2551	ATGCCCTATGGCTGCTCTTTAGACCATGTCCGGGAACCCGGAGCCCTGGCTCCCAG	2610
Db	1471	TACTGCTACCTGCAAGGGTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGGAA	1530	QY	653		653
QY	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480	Db	2611	GACCTGCTGAAGTGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGCGG	2670
Db	1531	CTGGGAGTGGACTGGCCCTCATCCACCATACACCCACTCTGCTTCTGTCACACGGTG	1590	QY	653		653
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	Db	2671	CTCGTACACAGGAGCTTGGCCGCTCGGAACGCTGCTGTCGAAGAGTCCCAACCATGTCAA	2730
Db	1591	CCCTGGGACAGCTCTTTCGGAACCCGACCAAGCTCTGCTCACACTGCCAACCGGCCA	1650	QY	653		653
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	Db	2731	ATTACAGACTTCGGGCTGGCTCGGCTGTGGAGATTGACGAGACAGAGTACCATGCAGAT	2790
Db	1651	GAGGACAGTGTGTGGCGAGGGCCCTGGCCCTGCCACCACTGTGGCCCGGAGGCACTGC	1710	QY	653		653
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540	Db	2791	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGAGTCCATTCCTCCGCGCGGTTTACC	2850
Db	1711	TGGGGTCCAGGGCCACCCAGTGTGTCACTGCAGCCAGTTCTTCTGGGGCCAGGAGTGC	1770	QY	653		653
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560	Db	2851	CACACAGATGATGTGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2910
Db	1771	GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGGAGTATGTGAATGCCAGGCATGT	1830	QY	653		653
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580	Db	2911	AAACCTTACGATGGGATCCACCGCGGAGATCCCTTGACCTGCTGGAAGGAGGAGCGG	2970
Db	1831	TTGCCGTGCCACCCCTGAGTGTACGCCCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG	1890				

```

QY 653 ----- 653
Db 2971 CTGCCCCAGCCCCCATCTGCACCATTTGATCTCTACATGATCATGGTCAAATGTTGGATG 3030
QY 653 ----- 653
Db 3031 ATTGACTCTGAATCTCGGCCAAGATTCGGGAGTGTGTCTGAATTTCTCCCGCATGGCC 3090
QY 654 ----- 654
Db 3091 AGGGACCCCCAGCGCTTTGTGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTG 3150
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACCTTCTACCGCTCACTGTGGAGACCATGACATGGGGAGCTTGGTGGATGCT 3210
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3211 GAGGAGTATCTGGTACCCAGCAGGGCTTCTTGTCTCAGACCTGCCCCGGGGCGCTGGG 3270
QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
Db 3271 GGCATGTCTCCACACAGCCAGCGCTCATCTACAGAGAGTGGCGGTGGGACCTGACA 3330
QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3331 CTAGGGCTGGAGCCCTCTGAAGAGAGAGGCCGCCAGGTCTCCACTGGCACCTCCGAGGG 3390
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3391 GCTGGCTCCGATGTATTGATGGTGACCTGGGAATGGGGGAGCAGCAAGGGCTGCAAGC 3450
QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3451 CTCCCCACACATGACCCCGCCCTCTACAGCGGTACAGTGAAGGCCACACAGTACCCCTG 3510
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3511 CCCCTCTGAGACTGATGGCTACGTTGCCCTGACCTGCAGCCGCCAGCCCTGAAATATGTG 3570
QY 805 AsnGlnProAspValArgProGlnProSerProArgGluGlyProLeuProAlaAla 824
Db 3571 AACCAGCCAGATGTTGGCCCCCAGCCCCCTTCGCCCGAGAGGGCCCTCTGCCTGTGCC 3630
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3631 CGACCTGCTGGTGCCTCTGGAAGGCCCAAGACTCTCTCCCGAGGAGAGATGGGGTC 3690
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3691 GTCAAAGACGTTTTTGGCTTTGGGGTGCCGTGGAGAACCCCGAGTACTTTGACACCCAG 3750
QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3751 GGAGGAGCTGCCCTCAGCCCCACCCCTCTCCCTGCTCAGCCAGCCAGCTTCGACAACTC 3810
QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3811 TATTACTGGGACAGGACCCACACAGAGGGGGCTCCACCCAGCACCTTCAAAGGACA 3870
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3871 CCTACGGCAGAGAACCCAGAGTACTTGGGTCTGGACCTGCCAGTG 3915

```

RESULT 13

```

US-10-101-510-124
; Sequence 124, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012

```

```

; CURRENT APPLICATION NUMBER: US/10/101.510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 124
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-124

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 12 Gaps: 1

SEQ6 (1-919) x US-10-101-510-124 (1-4530)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProGlyAla 20
Db 151 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCGCCCTCTTGGCCCCCGGAGCC 210
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db* 211 GCAGAGCACCACAGTGTGCACCGGCACAGACATGAAGCTGGGGCTCCCTGCCAGTCCCGAG 270
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60
Db 271 ACCACCTTGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAACCTG 330
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAATCTACCTACCTGCGCCACCAATGCCAGGCTGTCTCTTCCTGCAGGATATCCAGGAGGTG 390
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACCTGTCTATCGCTCACACCAAGTAGGCGAGGTCCACCTGCAGAGAGCTCGGG 450
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATGTGGGAGGCCACCGCTCTTTGAGGACAACTATGCCCTGGCGGTGCTAGACAAATGA 510
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAATACCAACCCCTGTCTCACAGGGGCTCCCGCAGGAGGCTCGGGAGCTG 570
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCGAAGCTCTACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAAACCCCGAG 630
QY 161 LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGTCTACCCAGCACAGATTTTGTGAAGSACATCTTCCACAGAACAACACAGCTGGCT 690
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCCGCTCTCGGGGCTGCCACCCCTGTCTCCGATGTGTAAG 750
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGCTGGGAGAGATTTCTGAGGATTTGTGAGAGCTTCAGAGCTGCAGCGCACTGTCTGT 810
QY 221 AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 811 GCCGGTGGCTGTCGCCCTGCAGAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGGCGGCTGCACGGGCCCAAGACACTCTGACTGCTGGCTGCTGCCCTGCCACTTCAACACC 930

```


QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 931 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACACACACACGTTTGAG 990
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 991 TCCATGCCCCAATCCCGAGGGCCGGTATACATTTCGGCGCCAGCTGTGTGACTCCCTGTCT 1050
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 1051 TACAACCTACCTTTACGGACGTGGATCCTGCACCCCTGTCTGCCCTCGCACACCCAA 1110
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 1111 GAGGTGACACAGAGAGTGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1170
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1171 GTGTGCTATGGTCTGGGCATGGAGCACCTTGCAGAGGTGAGGGCATACCAAGTGCCCAAT 1230
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 Db 1231 ATCCAGAGTTTGTGGCTGCAAGAGATCTTTGGAGCCTGGCATTTCTGCCGGAGAGC 1290
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400
 Db 1291 TTTGATGGGACCCAGCCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCCAAGTGT 1350
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 Db 1351 GAGACTCTGGAGAGATCAGAGTTACCTATATACATCTCAGCATGGCCGAGACAGCTGCCT 1410
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1411 GACCTCAGCGTCTTCAGAACCTGCAGTAATCCGGGGAGCAATTCTGCACAAATGCGGCC 1470
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1471 TACTCGCTCACCTGCAGGGCTGGGCATCAGCTGGCTGGCTGGCTGCCTCACTAGGGAA 1530
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1531 CTGGGCGATGGAGCTGGCCCTCATCCACCATACACCCACTCTGTCTGTGCACACGGTG 1590
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1591 CCTTGGGACCACTCTTTCGGAACCCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1650
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 Db 1651 GAGGACGAGTGTGGGCGAGGGCTGGCTGCCACCACTGTGGCCCGAGGGCAGCTGC 1710
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 Db 1711 TGGGGTCCAGGGCCACCAGTGTGTCAACTGCAGCCAGTTCTCTTGGGGCCAGGAGTGC 1770
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 Db 1771 GTGGAGGAATGCCGAGTCTGAGGGGCTCCCGAGGGAGTATGTAATGCCAGGCACCTGT 1830
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1831 TTGCCGTGCCACCTCAGTGTGAGCCCAAGATGGCTCAGTCACCTGTTTGGACCGGAG 1890
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 Db 1891 GCTACCACTGTGTGGCCCTGCCCCACTATAAGACCCCTCCCTTCTGCGTGGCCGCTGC 1950
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 Db 1951 CCCAGCGGTGAAACCTGACCTCTCTTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010

QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 Db 2011 GCGCATGCCAGCCTTGCCCATCAACTCACTGACCCACTCTCTGTGTGACCTGTGATGACAA 2070
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 Db 2071 GGCTCCCCCGGAGCAGAGAGCCAGCCCTCTGACGTCCCTGCTCTGCGGTGTGTGGC 2130
 QY 653 653 653
 Db 2131 ATCTCTGCTGGTGTGTGCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAG 2190
 QY 653 653 653
 Db 2191 AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGTGTGAGCCGCTG 2250
 QY 653 653 653
 Db 2251 ACACCTAGCGGAGCGATGCCCAACCGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTG 2310
 QY 653 653 653
 Db 2311 AGGAAGGTCAAGGTCTTGGATCTGGCGCTTTTGGCACACAGTCTACAGGGCATCTGGATC 2370
 QY 653 653 653
 Db 2371 CCTGATGGGAGAATGTGAAAATTCAGTGGCCATCAAAAGTGTGAGGGAAAAACACATCC 2430
 QY 653 653 653
 Db 2431 CCCAAAGCCAAACAAAGAAATCTTAGACCAAGCATACGTGATGGTGTGGGTGCCCTCCCA 2490
 QY 653 653 653
 Db 2491 TATGTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGTGACACAGCT 2550
 QY 653 653 653
 Db 2551 ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCCGGGACGCTGGGCTCCACG 2610
 QY 653 653 653
 Db 2611 GACCTGCTGAACCTGTGTATGCAGATTGCCAAGGGGATGAGTACCTGGAGGATGTGGC 2670
 QY 653 653 653
 Db 2671 CTCGTACACAGGAGCTTGGCCCTCGGAACGTGTGGTCAAGAGTCCCACCATGTCAAA 2730
 QY 653 653 653
 Db 2731 ATTACAGACTTCGGGCTGGCTGGCTGTGGACATTGACGAGACAGAGTACCATGCAGAT 2790
 QY 653 653 653
 Db 2791 GGGGCAAGGTCCCATCAAGTGGATGGCGTGGAGTCCATTCTCCGCGCGCGGTTCACC 2850
 QY 653 653 653
 Db 2851 CACCAGACTGATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2910
 QY 653 653 653
 Db 2911 AAACCTTAGATGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGGAGCGG 2970
 QY 653 653 653
 Db 2971 CTGCCCCAGCCCCCATCTGCACCATGATGTGTATATGATCATGTTGTTGATGATG 3030
 QY 653 653 653
 Db 3031 ATTGACTCTGAATGTGGCCCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC 3090
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664

[illegible]

691	CTCACACTGATAGACACCAACCGCTCTGGGCGCTGCCACCGCTGTCTCCGATGTGTAAAG	750
201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
751	GGCTCCCGCTCTGGGGAGAGATTCTGAGGATTGTCAGAGCTTGACGGCACTGCTCTGT	810
221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
811	GCGCGTGGCTGTGCCCGCTGCAAGGGGCGCACTGGCCCACTGCTGTGTCATGAGCAGTGT	870
241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
871	GCTGCCGCTGCAGGGGCGCCCAACGACTCTGACTGCGCTGGCGCTGCCTCCACTTCAACCAC	930
261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
931	AGTGGCATCTGTGAGCTGCATGCGCCAGCCCTGGTCACTACAACACAGACAGCTTTTGG	990

691	CTCACACTGATAGACACCAACCGCTCTGGGCGCTGCCACCGCTGTCTCCGATGTGTAAAG	750
201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
751	GGCTCCCGCTCTGGGGAGAGATTCTGAGGATTGTCAGAGCTTGACGGCACTGCTCTGT	810
221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
811	GCGCGTGGCTGTGCCCGCTGCAAGGGGCGCACTGGCCCACTGCTGTGTCATGAGCAGTGT	870
241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
871	GCTGCCGCTGCAGGGGCGCCCAACGACTCTGACTGCGCTGGCGCTGCCTCCACTTCAACCAC	930
261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
931	AGTGGCATCTGTGAGCTGCATGCGCCAGCCCTGGTCACTACAACACAGACAGCTTTTGG	990

QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	991	TCATATGCCAATCCGAGGCGCGTATACATTCGGCGCCAGCTGTGTGACTGCCTGCC	1050
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1051	TACAACCTACCTTTCTACGACGCGGGATCCTGCACCCCTGCTGCGCCCTGCACAAACCA	1110
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1111	GAGGTGACAGCAGGATGGAACACACAGCGGTGTGAGAAGTGCAGCAGCCCTGTGCCGA	1170
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1171	GTGTGTATGCTGGGCATGGAGCAGCTTCCGAGAGGTGAGGCGACTTACCACTGCCAAT	1230
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1231	ATCCAGAGTTTCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGGCGGAGAGC	1290
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1291	TTTGATGGGACCCAGCCTCCACACTGCCCGCTCCAGCCAGAGCTCCCAAGTGT	1350
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1351	GAGACTCTGGAAGAGATCACAGTTTACCTATACATCTCAGCATGTCGCGGACAGCCTGCCT	1410
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1411	GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGAGCGAATTCGCACAAATGCGCC	1470
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1471	TACTCGTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCAGTACAGGAA	1530
QY	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1531	CTGGGAGTGGAGTGGCCCTCATCCACCATAACACCACCTCTGCTTCTGTCACACGGTG	1590
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1591	CCCTGGACACAGCTCTTTCCGAACCCGACCAAGCTCTGCTCCACACTGCAACCGGCCA	1650
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1651	GAGGACGAGTGTGGGCGAGGCGCTGGCTGCCACACAGCTGTGCGCCGCGGAGCACTGC	1710
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1711	TGGGGTCCAGGGCCACCCAGTGTGTCACTGCAGCGCAGTTCTTCCGGGCGCAGGAGTGC	1770
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1771	GTGGAGAATGCCAGTACTGCAAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1830
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1831	TTGCCGTGCCACCTCAGTGTCAAGCCCGAGATGGCTCAGTACCTGTTTGGACCGGAG	1890
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1891	GCTGACAGTGTGGCCCTGTGCCACTATAGGACCCCTCCCTTCTGCGCTGGCCGCTGC	1950
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1951	CCAGCGGTGTGAACCTGACCTCTCTACATGCCATCCCTTCTGGAAGTTTCCAGATGAGGAG	2010
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Db	2011	GGCGCATGCCAGCTTGCCCATCACTGCACCCACTCTCTGTGTGACCTGGATGACAAG	2070
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2071	GGTGCCCGCGAGCAGAGCCCTCTGACGTCCATCGTCTGCGTGGTGGC	2130
QY	653	-----	653
Db	2131	ATTCTCTCGTCTGCTGCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGCAGCAG	2190
QY	653	-----	653
Db	2191	AAGATCCGGAAGTACACGATGCGGAGACTGCTCAGAGAAACGAGAGTGGTGGAGCCGCTG	2250
QY	653	-----	653
Db	2251	ACACCTAGCGAGCGATGCCCAACAGCGCGAGATCGGATCCTGAAAGAGACGAGCTG	2310
QY	653	-----	653
Db	2311	AGGAAGGTGAAGTGTCTGGATCTGCGCTTTTGGCACAGTCTACAAGGGCATCTCGATC	2370
QY	653	-----	653
Db	2371	CCTGATGGGAGNATGTGAAATTTCCAGTGGCCATCAAAGTGTGAGGGAAACACATCC	2430
QY	653	-----	653
Db	2431	CCAAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGGTGGTGTGGCTCCCA	2490
QY	653	-----	653
Db	2491	TATGCTCCGCTTCTGGGCATCTGCCCTGACATCCACGGTGCAGTGTGTGACACAGCTT	2550
QY	653	-----	653
Db	2551	ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCGCGGACGCTGGCTCCAG	2610
QY	653	-----	653
Db	2611	GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGATGAGTACCTGGAGGATGTCCGG	2670
QY	653	-----	653
Db	2671	CTCGTACAGGAGCTTGGCGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAAA	2730
QY	653	-----	653
Db	2731	ATTACAGACTTCGGGCTGCTCGCTGCTGGCATTTGACGAGACAGATACCATGCAGAT	2790
QY	653	-----	653
Db	2791	GGGGCAAGGTGCCCATCAAGTGGATGGGCTGGAGTCCATTCTCCGCGCGGCTTACC	2850
QY	653	-----	653
Db	2851	CACCAGATGATGTGTGGAGTTATGCTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2910
QY	653	-----	653
Db	2911	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG	2970
QY	653	-----	653
Db	2971	CTGCCCCAGCCCCCATCTGCACCATTTGATGCTTACATGATCATGTCATAATGTTGGATG	3030
QY	653	-----	653
Db	3031	ATTGACTCTGAATGTCCGCCAAGATTCGGGAGTTGGTGTCTGAATTCCTCCCATGGCC	3090
QY	654	-----	654
Db	3091	AGGACCCCCAGCGCTTTGTGGTCTCAGAAATGAGACTTTGGGCCAGCCAGCTCCCTTG	3150
QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspValAspAla	684
Db	665	-----	684

Db 3151 GACAGCACTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGTGGATGCT 3210
Qy 685 GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly 704
Db 3211 GAGAGTATCTGTTACCCACAGGGCTTCTTCTCCAGACCTGCCCGGGCGCTGGG 3270
Qy 705 GlyMetValHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
Db 3271 GGCATGGTCCACACAGCAGCCAGCTCATCTACAGGAGTGGCGGTGGGACCTGACA 3330
Qy 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3331 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCAGGTCTCCACCTGGCACCCCTCCGAAGG 3390
Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3391 GCTGGCTCCGATGATTGATGGTGACCTGGGAATGGGGAGCGCCAGGGGCTCAAGC 3450
Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3451 CTCGCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAGGAGCCACAGTACCCCTG 3510
Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3511 CCTCTGAGACTGATGGCTACGTTGCCCTGACCTGCAGCCCCAGCCCTGAATATGTG 3570
Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3571 AACCAGCAGATGTTGGCCCCAGCCCTTCGCCCGGAGAGGGCCCTCTGCTGCTGCC 3630
Qy 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3631 CGACCTGCTGTGGCCTCTCTGAAAGGGCCCAAGACTCTCTCCCGAGGGAAGATGGGTC 3690
Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3691 GTCAAGACGTTTGTGGCTTTGGGGGTCCGCTGGAGAACCCCGAGTACTTGACACCCAG 3750
Qy 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3751 GGAGGAGTGCCCTCAGCCCCACCTCTCTCTGCTTACGCCAGCCCTTCGACACCTC 3810
Qy 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3811 TATTACTGGGACAGGACCCACAGCGGGGGCTCCACCCAGCACCTTCAAGGGACA 3870
Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3871 CCTAGCGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG 3915

RESULT 15

US-10-177-293-125
; Sequence 125, Application US/10177293
; Publication No. US2003012128A1

GENERAL INFORMATION:

; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.

; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 125
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-177-293-125

Alignment Scores:

Pred. No.: 0 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 14 Gaps: 1

SEQ6 (1-919) x US-10-177-293-125 (1-4530)

Qy 1 MetGluLeuAlaLeuCysArgTyrGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGGAGCTGGCGGCTTGTGCCCTGGGGCTCTCTCGCCCTTTGCCCGCGAGCC 210
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCGAGCACCAAGTGTCCCGGCACAGACATGAAGCTCGGCTCCCTGCCAGTCCCGAG 270
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTCAGGGAACCTG 330
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAACCTACCTTACCTGCCCAATGCCAGGCTGTCTCTTCTTCCAGAGATATCCAGGAGTG 390
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACCTGTCTATCGCTCACACCAAGTAGGAGGGTCCACACTGCAGAGGCTCGG 450
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATGTGGGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCGGTGCTAGACAATGA 510
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAATACCACTCCCTGTCCAGGGGCTCCCGAGGAGGCTCGCGGAGCTG 570
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCGAG 630
Qy 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db 631 CTCTGCTACCCAGGACACGATTTTGTGAAGAGACATCTTCCACAGACACACAGCTGGCT 690
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
|||||

Dd	691	CTCACACTGATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGTAAG	750
Qy	201	GlySerArgCysTrpGluSerSerGluAspCysInserLeuThrArgThrValCys	220
Dd	751	GGCTCCCGCTGCTGGGAGAGAGTCTGAGGATGTTCAGAGCGCTGACGCGCACTGTCTGT	810
Qy	221	AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Dd	811	GCCGGTGGCTGTCCCGCTGCAGAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT	870
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Dd	871	GCTGCCGGCTGCACGGGGCCCAAGCACTCTGACTGCTGGCCCTGCCCTCAACCCAC	930
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Dd	931	AGTGGCATCTGTGAGCTGACTGCCAGCGCTGGTCACTACACACACACACAGCTTTGAG	990
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Dd	991	TCCATGCCAAATCCCGAGGGCCGGTATACATTCCGGCCAGCTGTGTGACTGCTGTCCC	1050
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Dd	1051	TACAACTACCTTTCTACGACGTGGGATCTTCACCCCTGCTGCCCCCTGCACAAACCAA	1110
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Dd	1111	GAGTGACAGCAGAGATGGAACACACAGCGGTGTGAGAAGTGCACGAAGCCCTGTGCCGA	1170
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Dd	1171	GTGTGTATGCTTGGCATGGAGCACTTGCAGAGAGTGAAGGCATTACCACTGCGCAAT	1230
Qy	361	IleGlnClnPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Dd	1231	ATCCAGAGTTTCTGGCTGCAGAGAAGATCTTTGGGAGGCTGGCATTTCTGCCGGAGAGC	1290
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe	400
Dd	1291	TTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1350
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Dd	1351	GAAGACTTGGAGAGATCAAGTTTACCTATACATCTCAGCATGGCGGACAGCTGCCT	1410
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Dd	1411	GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACAAATGGCGC	1470
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Dd	1471	TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGAA	1530
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Dd	1531	CTGGGAGAGTGGCTCATCTCCACCACTGCAAGTAATCCGGGGAGCAATCTGCACAAATGGCGC	1590
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Dd	1591	CCCTGGGACCAAGCTCTTTCCGAACCCGACCAAGCTCTGCTCCACTGCCAACCGGCCA	1650
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Dd	1651	GAGGACAGTGTGTGGCGAGGGCTTGGCTGCCACAGCTGTGCCCGGAGGGCACTGC	1710
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Dd	1711	TGGGGTCCAGGGCCCAAGTGTGCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1770
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Dd	1771	GTGGAGGAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1830
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Dd	1831	TTGCCGTGCCACCTGAGTGTGAGCCCCAGAAATGGCTCAGTACCTGTGTTTGGACCGGAG	1890
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Dd	1891	GCTGACCACTGTGTGGCCTGTGCCCACTATAAGGACCCCTCCCTTCTGCTGCGCCGCTGC	1950
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Dd	1951	CCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG	2010
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Dd	2011	GGGCATGCCAGCTTGGCCCCATCACTGCACCCACTCTCTGTGTGACCTGGATGACAAG	2070
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Dd	2071	GGCTGCCCGCGCAGCAGAGAGCCAGCCCTCTGACCTCCATCTCTCTGCGTGTGGC	2130
Qy	653	-----	653
Dd	2131	ATTCTGCTGCTGCTGCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGACAG	2190
Qy	653	-----	653
Dd	2191	AGATCCGGAAGTACAGATGCGGACACTGCTCCAGAAACGGAGCTGGTGGAGCCGCTG	2250
Qy	653	-----	653
Dd	2251	ACACCTAGCGGAGCGATGCCAACAGGCGCAGATCGCGATCCTGAAAGAGAGCGGAGCTG	2310
Qy	653	-----	653
Dd	2311	AGAAAGTGAAGTGTGATGCTGGCGCTTTGGCACAGTCTACAAGGGCATCTGGATC	2370
Qy	653	-----	653
Dd	2371	CCTGATGGGAGAACTGAAATTCAGTCAGTGCCATCAAAGTGTGAGGGAACACATCC	2430
Qy	653	-----	653
Dd	2431	CCCAAAGCAACAAGAAATCTTAGACGAAGCATACGTGATGGCTGTGGCTCCCCA	2490
Qy	653	-----	653
Dd	2491	TATGCTCCCGCTTCTGGGCATCTGCTTGACATCCACGGTGCAGCTGGTGACACAGCTT	2550
Qy	653	-----	653
Dd	2551	ATGCCCTATGCTGCTCTTAGACCATGTCCGGGAAACCCGGGAGCGCTGGCTCCCCAG	2610
Qy	653	-----	653
Dd	2611	GACCTGCTGAAGTGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGAGGATGTGCGG	2670
Qy	653	-----	653
Dd	2671	CTCGTACACAGGAGCTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA	2730
Qy	653	-----	653
Dd	2731	ATTACAGACTTCGGGCTGGCTGGCTGGACATTGACGAGACAGATACCATGCAGAT	2790
Qy	653	-----	653
Dd	2791	GGGGCAAGGTGCCCATCAAGTGTGAGTGGCGCTGAGTCCATTCTCCGCGCGGCTTACC	2850
Qy	653	-----	653
Dd	2851	CACCAGTGTGTGGAGTTATGTGTGACTGTGCGGAGCTGATGACTTTTGGGGCC	2910

```

QY 653 ----- 653
Db 2911 AAACCTTACGATGGGATCCAGCCGGGAGATCCTGACCTGCTGGAAGGGGAGCGG 2970
QY 653 ----- 653
Db 2971 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAAATGTTGGATG 3030
QY 653 ----- 653
Db 3031 ATTGACTCTGAATGTCGCCAAGATTCGGGAGTTGGTGTCTGAATTCCTCCGCATGGCC 3090
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 3091 AGGGACCCCGAGCGCTTTGTGTCATCCAGAATGAGGACTTGGGCCAGCCAGCTCCCTTG 3150
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGATGCT 3210
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyValaGly 704
Db 3211 GAGGAGTATCTGGTACCCAGCAGGCTTCTCTCTCCAGACCTGCCCGGGCGCTGGG 3270
QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
Db 3271 GGCATGGTCCACACAGCAGCCAGCTCATCTACAGGAGTGGCGGTGGGACCTGACA 3330
QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3331 CTAGGGCTGGAGCCCTCTGAAGAGAGGCCCCCAGGCTCCACTGGCACCTCCCGAAGGG 3390
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3391 GCTGGCTCCGATGATTTGATGGTGACCTGGGAATGGGGGAGCCAGGGGCTGCCAAGC 3450
QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3451 CTCGCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAGGAGGCCCCACAGTACCCCTG 3510
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3511 CCCTCTGAGACTGATGGTACGTTGCCCCCTGACCTGCAGCCCCCAGCCCTGAATATGTG 3570
QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3571 AACCAGCCAGATGTTGGCCCCCAGCCCTTCGCCCGAGAGGGCCCTCTGCCCTGCTGCC 3630
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3631 CGACTGTGTGTCACCTCTGGAAGGGCCCAAGACTCTCTCCCGAGGGAAGATGGGGTC 3690
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3691 GTCAAAGACGTTTGTGCTTTGGGGGTGCCCTGGAGAACCCCGAGTACTTGACACCCAG 3750
QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3751 GGAGGAGTGCCCTCTACGCCCCAGCCCTCTCTGCTTCAGCCCGAGCCCTTCGACAACTC 3810
QY 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3811 TATTACTGGGACGAGGACCCAGAGGGGGGCTCCACCCAGCACCTTCAAAGGGACA 3870
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3871 CCTACGGCAGAGAACCACAGTACCTGGGTCTGGACGTGCCAGTG 3915

```

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: October 15, 2003, 22:55:19 : Search time 232.963 Seconds
(without alignments)
1741.185 Million cell updates/sec

Title: SE06
Perfect score: 5078
Sequence: 1 MELAALCRWGLLLALLPPGA.....TFKGTPTAENPEYLGIDVPV 919

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPRO.spool/p/HOLLERAN80/runat_15102003_131913_20570/app_query.fasta_1.4685
-DB=Issued_Patents_NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN80 -CGN_1_1_403 @runat_15102003_131913_20570 -NCPU=6 -ICPU=3
-DEV_MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-NO_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	4900	96.5 3768 2	US-08-625-101-1
2	4900	96.5 3768 2	US-08-356-786-1
3	4900	96.5 4473 2	US-09-048-804-1
4	4900	96.5 4473 3	US-09-056-105-26
5	4892	96.3 4530 1	US-08-229-515A-9
6	4892	96.3 4530 1	US-08-645-865-9
7	4892	96.3 4530 4	US-09-167-322-4
8	4892	96.3 4530 4	US-09-527-487-1
9	4892	96.3 4530 4	US-09-877-177A-11
10	4134	81.4 3955 1	US-08-229-515A-14
11	4134	81.4 3955 1	US-08-645-865-14
12	3632	71.5 2385 2	US-09-146-283-3

13	3632	71.5	2385	3	US-08-579-823A-3	Sequence 3, Appli
14	3632	71.5	2385	3	US-09-344-195-3	Sequence 3, Appli
15	3471	68.4	1872	3	US-08-422-108-2	Sequence 2, Appli
16	3471	68.4	1872	4	US-08-422-734-2	Sequence 2, Appli
17	1645	32.4	5532	2	US-08-475-035-3	Sequence 3, Appli
18	1645	32.4	5532	4	US-09-676-610B-17	Sequence 17, Appli
19	1528.5	30.1	5484	3	US-09-632-580A-3	Sequence 3, Appli
20	1528.5	30.1	5501	1	US-08-484-438-1	Sequence 1, Appli
21	1451	28.6	5555	1	US-08-484-438-3	Sequence 3, Appli
22	1435	28.3	4905	1	US-07-978-895-3	Sequence 3, Appli
23	1435	28.3	4905	1	US-08-473-119-3	Sequence 3, Appli
24	1435	28.3	4905	2	US-08-475-352-3	Sequence 3, Appli
25	1435	28.3	4975	3	US-09-630-706-3	Sequence 3, Appli
26	1311	25.8	4545	6	5183884-3	Patent No. 5183884
27	1032	20.3	1958	4	US-09-570-454-1	Sequence 1, Appli
28	1032	20.3	1958	4	US-09-867-521-1	Sequence 1, Appli
29	942	18.6	1593	4	US-09-676-610B-25	Sequence 25, Appli
30	942	18.6	1868	1	US-08-658-883B-1	Sequence 1, Appli
31	942	18.6	1868	4	US-09-676-610B-26	Sequence 26, Appli
32	493	9.7	322	1	US-08-421-356-1	Sequence 1, Appli
33	493	9.7	322	4	US-09-046-783-1	Sequence 1, Appli
34	381	7.5	4149	2	US-08-737-715-1	Sequence 1, Appli
35	376	7.4	2533	3	US-09-149-922-5	Sequence 5, Appli
36	335	6.6	4989	2	US-08-625-819-1	Sequence 1, Appli
37	333	6.6	4975	2	US-08-249-687C-1	Sequence 1, Appli
38	333	6.6	4989	3	US-08-666-392A-3	Sequence 3, Appli
39	333	6.6	4989	3	US-08-755-558-4	Sequence 4, Appli
40	333	6.6	4989	3	US-08-746-559A-1	Sequence 1, Appli
41	333	6.6	4989	3	US-08-880-313A-9	Sequence 9, Appli
42	333	6.6	4989	3	US-09-199-926-3	Sequence 3, Appli
43	333	6.6	4989	4	US-09-389-855A-9	Sequence 9, Appli
44	333	6.6	4989	4	US-09-668-822-9	Sequence 9, Appli
45	328	6.5	4993	3	US-08-746-559A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-625-101-1
Sequence 1, Application US/08625101
Patent No. 5869445
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/625,101
FILING DATE: 01-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C7
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3768 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..3765
 ; US-08-625-101-1

Alignment Scores:

Pred. No.: 0 Length: 3768
 Score: 4900.00 Matches: 919
 Percent Similarity: 73.23% Conservative: 0
 Best Local Similarity: 73.23% Mismatches: 0
 Query Match: 96.49% Indels: 336
 DB: 2 Gaps: 1

SE06 (1-919) x US-08-625-101-1 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProProGlyAla 20
 |||||
 Db 1 ATGGAGCTGGCGCCTGTGGCGCTGGGGGCTCCTCCTCGCCCTCTTGCCCCCGGAGCC 60
 |||||
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 |||||
 Db 61 GCGAGCACCCCAAGTGTGCACCGGCACAGACATGAAGCTCGGGCTCCTGCCAGTCCCGAG 120
 |||||
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 |||||
 Db 121 ACCCACTGGACATGTCCGCCACCTCTACCAAGGCTGCCAGGTGGTGAGGGAACCTG 180
 |||||
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 |||||
 Db 181 GAACCTACCTACCTGCCACCAATGCCAGCCTGTCTCTCGAGGATATCAGAGGTG 240
 |||||
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 |||||
 Db 241 CAGGGCTAGCTGTCTATCGCTCACACCAAGTGCAGCAGGTCCCACTGCAGAGGTCCGG 300
 |||||
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 |||||
 Db 301 ATTGTCCGAGGACCCAGCTCTTTGAGGACACTATGCCCTCGCCGTGTAGCAATGGA 360
 |||||
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyLeuArgGluLeu 140
 |||||
 Db 361 GACCCCTGAACAATACCACTCTGCACAGGGCTCCCGAGGAGCTCGGGAGCTG 420
 |||||
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 |||||
 Db 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCAG 480
 |||||
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 |||||
 Db 481 CTCGTCTACAGACACAGATTTGTGAAGAGGACATCTCCACAAGAACCAACAGGTGGCT 540
 |||||
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 |||||
 Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTAG 600
 |||||
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 |||||
 Db 601 GGCTCCCGCTGTGGGGAGAGAGTTCTGAGGATTGTGAGAGCTGTCAGAGCTGACGGCCTGCTCTGT 660
 |||||
 QY 221 AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
 |||||
 Db 661 GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 |||||
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 |||||
 Db 721 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 |||||
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 |||||

Db 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTGCACTACAAACACACACACGTTTGAG 840
 |||||
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 |||||
 Db 841 TCCATGCCCATCCCGAGGCGCGGTATACATTCGGGCGCAGCTGTGTACTGCTGTCTCC 900
 |||||
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 |||||
 Db 901 TACAACCTACCTTTCTACGGACGTGGGATCCTGCACCTCGTCTGCCCTGTCACAAACCAA 960
 |||||
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 |||||
 Db 961 GAGGTGACACGACAGAGATGGAACACAGCGGTGTGAGAAGTGAGCAAGCCCTGTGCGCGA 1020
 |||||
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 |||||
 Db 1021 GTGTGCTATGTGCTGGGCATGGAGCACTTGCAGAGGTGAGGCGAGTTACCAAGTGCAT 1080
 |||||
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 |||||
 Db 1081 ATCCAGGAGTTTGTCTGGCTGCAAGAAGATCTTTGGAGGCTGSCATTTCTGCCGAGAGC 1140
 |||||
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 |||||
 Db 1141 TTTGATGGGGACCCAGCCTCCAAACACTGCCCGCTCCAGCCAGAGCAGTCCAAGTGT 1200
 |||||
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 |||||
 Db 1201 GAGACTCTGGAAGAGATCACAGTTACCTATATACATCTCAGCATGGCCGACAGCTGCCT 1260
 |||||
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 |||||
 Db 1261 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGAGCAATCTGACAAATGGCGCC 1320
 |||||
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 |||||
 Db 1321 TACTCGCTGACCTGCAGGGCTGGGCATCAGCTGGCTGGGCTGGCTGCTGCTGCTGCTGCTGCT 1380
 |||||
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 |||||
 Db 1381 CTGGGAGTGGACTGGCCCTCATCCACCATAACCCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 |||||
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 |||||
 Db 1441 CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
 |||||
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 |||||
 Db 1501 GAGGACGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGCTGCCCGCCGAGGGCACTGC 1560
 |||||
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 |||||
 Db 1561 TGGGGTCCAGGGCCCCCAGCGTGTGTCAACTGCAGCAGTCTCCTCTCGGGGGCCAGAGTGC 1620
 |||||
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 |||||
 Db 1621 GTGGAGGAATCCGAGTACTGCAGGGGCTCCCGCAGGAGATATGTGAATGCCAGGCACTGT 1680
 |||||
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 |||||
 Db 1681 TTTGCCGTGCCACCTGAGTGTGAGCCCCAGAAATGCTCAGTAGCCTGTGTTTGGACCGGAG 1740
 |||||
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 |||||
 Db 1741 GCTGACCACTGTGTGGCTGTGCCCACTATAGGACCTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 |||||
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 |||||
 Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
 |||||
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 |||||


```

: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA: US/08/356,786
: APPLICATION NUMBER: US/08/356,786
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/831,967
: FILING DATE: 06-FEB-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Pitcher, Edmund R.
: REGISTRATION NUMBER: 27,829
: REFERENCE/DOCKET NUMBER: CRP-053
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 248-7000
: TELEFAX: (617) 248-7100
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3768 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..3768
: OTHER INFORMATION: /note= "cerB-b2"
: US-08-356-786-1

```

Alignment Scores:

```

Pred. No.: 0 Length: 3768
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 2 Gaps: 1

```

SEQ6 (1-919) x US-08-356-786-1 (1-3768)

```

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProGlyAla 20
DB 1 ATGGAGCTGGCGGCTGTGCCGCTGGGGCTCTCTCGCCCTCTTGGCCCGCGAGCC 60
QY 21 AlaserThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
DB 61 GCAGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
DB 121 ACCCAGCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
DB 181 GAATCAGCTTACCTGCTCCACCAATGCGGCTGCTCTCTCTGAGGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
DB 241 CAGGGTACGTGCTATCGCTCAACCAAGTAGGAGGAGGTCCCATCTGCAGAGGCTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
DB 301 ATTGTGGAGGACCCAGCTCTTTTGGAGCAACTATGCTCCCTGCCCTGTAGACAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
DB 361 GACCCGTGAACAATAACACCCCTGTGCAGAGGGCCCTCCCGAGAGGCTGCGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGln 160
DB 421 CAGCTTCGAGGCTCAGAGATCTTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCCAG 480

```

```

QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
DB 481 CTCTGCTACCAGGACACGATTTTGTGAAGGACATCTCCACAAGAACAACACGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
DB 541 CTCACACTGATAGACACCAACCCCTCTCGGGCCCTGCCACCCCTGTCTCCGATGTGTAA 600
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
DB 601 GGCTCCCGCTCTGGGGAGAGAGTTCTGAGATTGTGAGAGCTGAGAGCTGAGCGCACTGTCTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
DB 661 GCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCACCTGACTGCTGCCATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 721 GCTGCCGGCTGCACGGGCCCAAGCACCTCTGACTGCTGCCCTGCCCTTCCACTTCAACAC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
DB 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGTGTCACCTACACACACACAGCATTGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
DB 841 TCCATGCCCAATCCCGAGGGCCGATATACATTCCGGCCGACGCTGTGACTGCTGTGCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
DB 901 TACAACCTAGCTTTCTACGGACGTGGGATCCTGCAACCTCTGCTGCCCTTGCACAACCAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 961 GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGAGCCCTGTGCCCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1021 GTGTGCTATGCTGGGATGGAGCACTTGCAGAGGTGAGGCGAGTTACAGTGGCAAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
DB 1081 ATCAGGAGTGTCTGGCTGCAAGAAGATCTTTTGGAGGCTGGCATTTTCTGCCGAGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
DB 1141 TTTGATGGGAGCCACCCCTCCAACACTGCCCGCTCCAGCCAGCAGCTCCAGGTGTTT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
DB 1201 GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCGGACAGCTTGCCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
DB 1261 GACCTCAGGCTCTTCCAGAACCTGCAAGTAATCCGGGAGCAATCTGCACAATGGCGCC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
DB 1321 TACTCCTGACCTGCAAGGGCTGGCATCAGCTGGCTGGGCTGGCTCCTCAGTGGGAA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
DB 1381 CTGGGAGTGGACTGGCCCTCATCCACCATAAACCCACCTCTGCTTCTGTCACAGCGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
DB 1441 CCTGGGACAGCTCTTTTGGAAACCCGACCAAGCTTCTGCTCCACTTGCACACCGGCCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
DB 1501 GAGGACAGTGTGTGGCGGAGGCTTGGCTGCCACCACTGCTGCGCCGAGGAGGACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540

```



```

RESULT 3
US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; TITLE OF INVENTION: HUMAN HER-2 EXPRESSION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748ris LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048,804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: No
US-09-048-804-1

Alignment Scores:
Pred. No.: 0 Length: 4473
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 2 Gaps: 1

SEQ6 (1-919) x US-09-048-804-1 (1-4473)
QY 1 MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGCTGGCGCCTTGCGCGCTGGGGCTCCCTCGCCCTCTTGCCCGCCGAGGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGACCCAGTGTGACCGGCACAGACATGAGCTGGGCTCCCTGCCAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuAlaHisAsnGlnValArgGlnGlyCysGlnValGlnGlyAsnLeu 60
Db 295 ACCACCTGGACATGCTCGCCACCTCTACAGGGCTGCCAGGTGGTGGTGGAGAAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 355 GAACCTACCTAGCTGCCACCAATGCCAGCCTGCTCTTCTCAGGATATCCAGAGGTG 414
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGGCTAGCTGCTATCGCTCACACACCAAGTGAAGGAGGCTCCCACTGAGAGGCTGCGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120

```

```

Db 475 ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACACTATGCCCTGCCGTGTAGACAATGGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGAACAATACCACTCCCTGTCAAGGGGCTCCCAAGAGGCTGCGGGAGCTG 594
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGAGGGGCTTGTATCAGGGGAACCCCAAG 654
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db 655 CTCTGCTACCAAGGACAGATTTTGTGGAAGGACATCTTCCACAAGAACCAACAGCTGGCT 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACACTGATAGACCAACCGCTCTCGGGCTGCCACCTGTCTCCGATGTGTAAG 774
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 775 GGCTCCCGCTGCTGGGAGAGAGATTCTGAGGATTGTGAGAGCTGACGGGCACTGTCGT 834
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnCys 240
Db 835 GCGCGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCACTGCTGCTGCTGCTGCTGCTGCTGCT 894
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 895 GCTCCCGCTGCACGGGGCCCCAAGCACTCTGACTGCCCTGGCTGCTGCTGCTGCTGCTGCTGCT 954
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 955 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGTGAGAGTGCAGAGCCCTGTGCCCA 1014
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 1015 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGCTGCTGCTGCTGCTGCTGCT 1074
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1075 TACAACTACCTTTCTACGGACCTGGGATCCTGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1135 GAGGTGACAGCAGAGAGATGGAACACAGCGGTGTGAGAGTGCAGAGCCCTGTGCCCA 1194
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1195 GTGTGCTATGTGTGGCATGGAGCACTTGCAGAGAGTGCAGGCGAGTTACCAAGTGCCTAT 1254
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1255 ATCCAGGAGTGTGTGGCTGCAAGAGATCTTTGGAGCTGGCATTTCTGCCGAGAGAG 1314
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1315 TTTGATGGGAGCCAGCCCTCCAACACTGCCCGCTCCAGCCAGCAGAGCAGCTCCAAGTGT 1374
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro 420
Db 1375 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGTGGCGGACAGCTGCCT 1434
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1435 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCCGGGGAGCAATTCGCACCAATGGCGCC 1494
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1495 TACTCGCTGACCTGCAGAGGCTGGGCATCAGCTGGCTGGGGCTGGCTGCTGCTGCTGCTGCTGCTGCT 1554
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480

```

Db	1555	CTGGCAGTGGACCTGGCCCTCATCCACATCAACACCACCTCTCTCTGGTGCACACGGTG	1614
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1615	CCCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1674
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1675	GAGGACGAGTGTGTGGCGAGGCGCTGGCCCTGCCACCAGCTGTGGCCGAGGCGACTGC	1734
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1735	TGGGGTCCAGGGCCCAACCCAGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC	1794
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1795	GTGGAGNAATGCCAGTACTGCAGGGGCTCCCCAGGAGTATGTGAATGCCAGGCACGTGT	1854
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1855	TTCGCGTGCCACCTGAGTGTGACGCCCAAGATGGCTCAGTGACCTGTTTTGGACCGGAG	1914
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrIlyAspProPropheCysValAlaArgCys	600
Db	1915	GCTGACCAAGTGTGGCGCTGCCCCACTATTAAGACCTCCCTTCCTGCGTGGCCCGCTGC	1974
Qy	601	ProSerGlyValIlyProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1975	CCCAGGGTGTGAACCTGACCTCTCTACATGCCCCATCTGGAAGTTTCCAGATGAGGAG	2034
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAsIlys	640
Db	2035	GGCGCATGCCAGCTTGCCCATCACTGCACCCACCTCCTGTGTGGACCTGGATGACAAG	2094
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2095	GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGC	2154
Qy	653	-----	653
Db	2155	ATTCTGCTGGTGTGGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGCACGAG	2214
Qy	653	-----	653
Db	2215	AAGATCCGGAAGTACACGATGCGGAGACTGTGTGCGAAGACGGAGCTGGTGGACCGCTG	2274
Qy	653	-----	653
Db	2275	ACACCTAGCGGAGCGATGCCCCAACCCAGGCGCAGATGCGGATCCTGTAAAGAGACGGAGCTG	2334
Qy	653	-----	653
Db	2335	AGGAAGGTGAAGGTGGTCTGGTCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2394
Qy	653	-----	653
Db	2395	CCTGATGGGAGAAATGTGAAAAATCCAGTGGCCATCAAAAGTGTGTAGGGAAAAACACATCC	2454
Qy	653	-----	653
Db	2455	CCCAAGCCACAAAGAAATCTTTAGACGAAGCATACGTGATGGCTGTGTGGGTCCCCA	2514
Qy	653	-----	653
Db	2515	TATGTCCTCCGCCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2574
Qy	653	-----	653
Db	2575	ATGCCCTATGGCTGCCCTTTAGACCAATGTCCGGGAAAAACCGCGACGCCCTGGGTCCACG	2634
Qy	653	-----	653
Db	2635	GACCTGCTGAAGTGTGTATGCAGATTGCCAAGGGGATGAGTACCTGGAGGATGTGCGG	2694
Qy	653	-----	653
Db	2695	CTCGTACACAGGGACTTGGCCGCTCGGAACGCTGCTGGTCAAGAGTCCCAACCATGTCAA	2754
Qy	653	-----	653
Db	2755	ATTACAGACTTCCGGCTGGCTCGGCTGTGCACATTTGACGAGACAGAGTACCATGCAGAT	2814
Qy	653	-----	653
Db	2815	GGGGCAAGTGGCCATCAAGTGAATGGCGCTGGAGTCCATTCTCCGCCGGCGGTTCACC	2874
Qy	653	-----	653
Db	2875	CACCAGAGTGAATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC	2934
Qy	653	-----	653
Db	2935	AAACCTTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAAAAGGGAGCGG	2994
Qy	653	-----	653
Db	2995	CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAAAATGTTGGATG	3054
Qy	653	-----	653
Db	3055	ATTGACTCTGAATGTCGGCCCAAGATTCCGGGAGTTGGTGTCTGAATTCCTCCGCATGGCC	3114
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	3115	AGGACCCCCAGCGCTTTTGGTCAATCCAGAATCAGGACTTGGGCCCGCCAGCTCCCTTG	3174
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3175	GACAGCACTTCTTACCCTCCTGCTGGAGACATGACATGGGGAGCTGTGTGGATGCT	3234
Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3235	GAGGAGTATCTGGTACCCAGCAGGGCTTCTTCTGTCCAGACCTGCCCGGGCGCTGGG	3294
Qy	705	GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr	724
Db	3295	GGCATGTCCACCACAGCCAGCAGCTCATCTACCAGGAGTGGCGGTGGGACCTGACA	3354
Qy	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	3355	CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCAGGTCTCCACTGGCACCTCCGAAGGG	3414
Qy	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnIns	764
Db	3415	GCTGGCTCCCATGTATTGTGATGGTACCTGGGAATGGGGCAGCAGGGCTGCCAAGC	3474
Qy	765	LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu	784
Db	3475	CTCCCCACATGACCCAGCCCTCTACAGCGGTACAGTGGAGGACCCACAGTACCCTG	3534
Qy	785	ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal	804
Db	3535	CCCTCTGAGACTGATGGCTACGTGGCCCTTACCCTGACCTGCAGCCCCCAGCTGAATATGT	3594
Qy	805	AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla	824
Db	3595	AACCGCCAGATGTTCCGCCCCAGCCCTTCCGCCCGAGAGGGCCCTCTGCTGCTGCC	3654
Qy	825	ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal	844
Db	3655	CGACCTGCTGGTCCACTCTGGAAGGCCCAAGACTCTCTCCCGCAGGAAAGAAATGGGTC	3714
Qy	845	ValIlyAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln	864
Db	3715	GTCAAAGACGCTTTTGCCTTTGGGGTGGCTGGAGAACCCCGAGTACTTTGACACCCAG	3774

Qy	865	GlyGlyAlaAlaProGlnProHisProProProAlaAlaPheSerProAlaPheAspAsnLeu	884
Db	3775	GGAGGAGCTGCCCTCAGCCCCACCCCTCCCTCCCTTCAGCCAGCGCTTCGACAACTC	3834
Qy	885	TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr	904
Db	3835	TATTACTGGGACCAGGAGCCACCAGAGCGGGGGCTCCACCCAGCACTTCAAAGGGACA	3894
Qy	905	ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal	919
Db	3895	CCTACGGCAGAGAACCCAGAGTACTGGTCTGGACGTGCCAGTG	3939
RESULT 4			
US-09-056-105-26			
; Sequence 26, Application US/09056105			
; Patent No. 6287569			
; GENERAL INFORMATION:			
; APPLICANT: KIPPS, THOMAS J.			
; APPLICANT: WU, YUNQI			
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR			
; FILE OF INVENTION: PROCESSING			
; FILE REFERENCE: 233/221			
; CURRENT APPLICATION NUMBER: US/09/056,105			
; CURRENT FILING DATE: 1998-04-06			
; EARLIER APPLICATION NUMBER: 60/043,467			
; EARLIER FILING DATE: 1997-04-10			
; NUMBER OF SEQ ID NOS: 35			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 26			
; LENGTH: 4473			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-056-105-26			
Alignment Scores:			
Pred. No.: 0 Length: 4473			
Score: 4900.00 Matches: 919			
Percent Similarity: 73.23% Conservative: 0			
Best Local Similarity: 73.23% Mismatches: 0			
Query Match: 96.49% Indels: 336			
DB: 3 Gaps: 1			
SEQ6 (1-919) x US-09-056-105-26 (1-4473)			
Qy	1	MetGluLeuAlaAlaLeuCysArgTyrGlyLeuLeuAlaLeuLeuProProGlyAla	20
Db	175	ATGGAGCTGGCGGCTGTGGCGTGGGGCTCTCTCTCGCCCTCTGCCCCCGGAGCC	234
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	235	CGGAGCACCCAAAGTGTCCACGGGCACAGACATGAAGCTGGGGCTCCCTGCCAGTCCCGAG	294
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
Db	295	ACCCACCTGGACATGCTCCGCCACCTCTACCGGGCTGCCAGGTGGTGCAGGAAACCTG	354
Qy	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	355	GAACTACCTACCTGGCCCAATGCCAGCTGTCTCTCTGCAGGATATCCAGAGGGGTG	414
Qy	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	415	CAGGGCTACGTGCTCATCGCTCAACCAAGTAGGAGGTCCTCCACATGCAGAGGCTCGGG	474
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	475	ATTGTGCGAGGACCCAGCTCTTTAGGACAACATATGCCCTGGCCGTCTAGACAATGGA	534
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	535	GACCCGTGAACAATACACCCCTGTCCAGGGGCGCTCCCGAGAGCCCTGCGGGAGCTG	594
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuLeuGlnArgAsnProGln	160

Db 1675 GAGGACGAGTGTGGGGCGAGGGGCTGGCCCTGCCACACAGCTGTGGCGCCGAGGCGACTGC 1734
 Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 Db 1735 TGGGGTCCAGGGCCACCCAGTGTCACTGCGAGCCAGTTCTTCGGGGCCAGAGTGC 1794
 Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaAraHisCys 560
 Db 1795 GTGGAGGAATGCCGAGTACTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACCTGT 1854
 Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1855 TTGCCGTGCCACCTGTGAGTGTGAGCCCGCAGAAATGGCTGAGTGTCTTTTGGACCGGAG 1914
 Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAraCys 600
 Db 1915 GCTGACCAGTGTGGGCTGTGCCCACTATTAAGGACCTCCCTTCTGGGTGGCCCGCTGC 1974
 Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProLeuThrLysPheProAspGluGlu 620
 Db 1975 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGTAAGTGTCCAGATGAGGAG 2034
 Qy 621 GlyAlaCysGlnProCysProLleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 Db 2035 GCGCATGCCAGCTGTGCCCATCACTGACACCCACCTCTGTGTGGACCTGGATGACAAG 2094
 Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 Db 2095 GGCTGCCCGCCGAGAGAGAGCCGCTCTGACGTCCATCTCTGCGGTGGTTGGC 2154
 Qy 653 2155 ATTCTGTGTCGTGGTCTTGGGGTGTGCTTTGGGATCCCTCATCAAGCAGCGCAGCAG 2214
 Qy 653 2215 AAGATCCGGAAATACACGATGCGGAGACTGCTGCGAGAAACGGAGCTGGTGGAGCCGCTG 2274
 Qy 653 2275 ACACCTAGCGAGCGATGCCCAACCCAGCGCAGATGCGGATCTCTGAAGACAGCGAGCTG 2334
 Qy 653 2335 AGGAAGTGAAGGTCTTGGATCTGGCTTTTGGCACAGTCTACAAGGCGATCTGGATC 2394
 Qy 653 2395 CCTGATGGGGAGAAATGTGAAAATTCAGTGGCCATCAAAAGTGTGAGGGAAACACATCC 2454
 Qy 653 2455 CCCAAGCCCAAGAANAATCTTAGACGAAGCATACGTGATGGCTGGTGGGTCCCA 2514
 Qy 653 2515 TATGCTCCCGCTTCTGGGCATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2574
 Qy 653 2575 ATGCCCTATGGCTGCCCTTTAGACCATGTCCGGGAAACCGCGGACGCCCTGGGCTCCAG 2634
 Qy 653 2635 GACCTGTGAACTGGTGTATGCAGATTTGCCAAGGGGATGAGCTACCTGGAGGATGTCGG 2694
 Qy 653 2695 CTCGTACACAGGGACTTGGCGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAA 2754
 Qy 653 2755 ATTACAGACTTCGGGCTGGCTGCTGACATTTGACGACATGACGACAGAGTACCATGCAGAT 2814

Qy 653 2815 GGGGCAAGGTGCCCATCAAGTGGATGCGCTGGAGTCCATTCTCCGCCGGGGTTTACC 2874
 Qy 653 2875 CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2934
 Qy 653 2935 AAACCTTACGATGGATGCCAGCCCGGAGATCCCTGACCTGCTGGAAAAAGGGGAGCGG 2994
 Qy 653 2995 CTGCCCCAGCCCCCTCTGCACCATTTGATGCTTACATGATCATGGTCAAAATGTTGGATG 3054
 Qy 653 3055 ATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC 3114
 Qy 654 3115 AGGACCCCCAGCCCTTTGTGGTCAATCCAGAAATGAGGACTTGGGCCCGCCAGCCCTTGTG 3174
 Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 Db 3175 GACAGACCTTCTACCGCTCACTGCTGGAGGACCATGACATGGGGACCTGCTGGTGGTGT 3234
 Qy 685 GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly 704
 Db 3235 GAGGAGTATCTGGTACCCAGCAGGGCTTCTTCTGTCAGACCTGCCCCGGCGCTGGG 3294
 Qy 705 GlyMetValHisArgHisArgSerSerSerThrArgSerGlyGlyGlyAspLeuThr 724
 Db 3295 GGCATGTGTCCACACAGCACCAGCAGCTCACTTACCAGGAGTGGCGTGGGAGCTTGACA 3354
 Qy 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 Db 3355 CTAGGGCTGGAGCCCTCTGAAGAGAGGCCCCCAGGCTTCCACTGGCACCTCCGAAGGG 3414
 Qy 745 AlaGlySerAspValPheAspGlyAspLeuMetGlyAlaAlaLysGlyLeuGlnSer 764
 Db 3415 GCTGGCTCCGATGATTTGATGTGACCTGGGAATGGGGCAGCACAAGGGCTGCAAGC 3474
 Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 Db 3475 CTCCCCACATGACCCAGCCCTTACAGCGGTACAGTGAAGGACCCACAGTACCCCTG 3534
 Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 Db 3535 CCTCTGAGACTGATGGCTAGCTTGGCCCTGACCTGACGCCCCCGCCCTGAATATGTG 3594
 Qy 805 AsnGlnProAspValArgProGlnProSerProArgGluGlyProLeuProAlaAla 824
 Db 3595 AACCCAGCAGATGTTGGCCCGCCAGCCCTTCGCCCGCAGAGGGGCTCTGCTGCTGCTG 3654
 Qy 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
 Db 3655 CGACCTGCTGGTGGCTCTGGAAGGCGCCCAAGACTCTCTCCCCGAGGAAATGGGGTCT 3714
 Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
 Db 3715 GTCAAAACGCTTTTGGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGGACACCCAG 3774
 Qy 865 GlyGlyAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3775 GGAGGAGTGGCCCTCAGCCCGCCCTCTCTGCTGCTTCCAGCCCGAGCTTTCGACAACTC 3834
 Qy 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
 Db 3835 TATTACTGGGACGAGGACCCAGGAGGGGGCTCCACCCAGCAGCTTTCAAAGGACA 3894


```
|||||
Db 3691 GTCAAGACGCTTTTGGCTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCGAG 3750
QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAsnLeu 884
Db 3751 GGAGGAGCTCCCTCAGCCCCACCCTCCCTCCCTTCAGCCCGAGCTTCGACACCTC 3810
QY 885 TyrTyrTyrAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3811 TATTACTGGACACGAGACCCACAGAGCGGGGGCTCCACCCAGACCTTCAAAAGGACA 3870
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3871 CCTACGACAGAGAACCAGAGTACCTGGGTCTGGAGCTGCCAGTG 3915

RESULT 6
US-08-645-865-9
; Sequence 9, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPlastic DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-645-865-9

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 1 Gaps: 1

SEQ6 (1-919) x US-08-645-865-9 (1-4530)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCCTCTCGCCCTCTTGGCCCGCGAGCC 210
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 CGGAGCACCAAGTGTGCACCGCCGACAGATGAAGTGGGCTCCTTGCACGAGCTCCCTGCCAGTCCGAG 270
```

```
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCCACTGGACATGTCTCCGCCACCTCTACCAAGGGTGCAGAGTGGTGGAGGAAACCTG 330
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAACTCACCTTACCTGCCCAATGCCAGCTGCTCTCTCTGCTGAGGATATCCAGGAGGTG 390
QY 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGTACGTGCTCATCGCTCAACCAAGTAGGAGGCTCCCACTGCAGAGGCTGCGG 450
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCCCTGCTGAGACATGGA 510
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAATACCACTCCCTGTCACAGGGGCTCCCAAGAGGCTGCGGAGCTG 570
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuLeuGlnArgAsnProGln 160
Db 571 CAGCTTCGAAGCCTCAGAGATCTTGAAGAGGAGGCTCTTGATCCAGCGGAAACCCCGAG 630
QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAla 180
Db 631 CTCTGCTACCAAGGACAGATTTTGTGGAAGGACATCTTCCCAAGAACACACAGCTGGCT 690
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG 750
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGCTGGGAGAGATTCTGAGGATTGTGAGAGCTGAGCGGCATCTGTGT 810
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 811 GCCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGCTGCTGCTGCTGCTGCTG 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGCCGGTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTG 930
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 931 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGCTCACCAACACAGACAGCTTTGAG 990
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 991 TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGCTGCTGCTG 1050
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1051 TACAACCTACTTCTACGGACGTGGGATCTCTGCACCTCTGCTGCTGCTGCTGCTGCTG 1110
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysSerLysProCysAlaArg 340
Db 1111 GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGAGAGCTGTGCGCGGA 1170
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1171 GTGTGCTATGCTGCGCATGGAGCACTTGCAGAGAGTGAGGCGAGTTACAGTGCCAAT 1230
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1231 ATCCAGAGAGTTTCTGGCTGCAAGAAGATCTTTTGGAGCGCTGCTGCTGCTGCTGCTG 1290
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1291 TTTGATGGGAGCCAGCCCTCCAACACTGCCCCCTCCAGCCAGAGCACTTCCCAAGTGTTT 1350
```

QY 401 GluThrLeuGluGluLeuThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
 DB 1351 GAGACTCTGGAGAGATCACAGGTACCTATATACATCTCAGCATCGCCGACAGCCTGCCT 1410
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1411 GACCTCAGCGTCTCCAGAACCTGCAAGTAATCGGGACGAATTCGCACAATGGCGCC 1470
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTyrLeuGlyLeuArgSerLeuArgGlu 460
 DB 1471 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTCGCTCACTGAGGGAA 1530
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1531 CTGGCAGTGGACTGGCCCTCATCCACCATACACCCACCTCTGCTTCTGTCACACAGGTG 1590
 QY 481 ProTyrAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1591 CCTGGGACCAAGCTCTTCGGANACCGGCACCAAGCTCTGCTCCACACTGGCCACCGGCCA 1650
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1651 GAGGACGAGTGTGGGGAGGCGCTGGCTGCCCTGCCACCAAGCTGCTGCCCGGAGGCACTGC 1710
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1711 TGGGCTCAGGGCCCAACCACTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1770
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 DB 1771 GTGGAGGAATGCCAGTACTCAGGGGCTCCCGAGGAGTATGTAATGCCAGGCACCTGT 1830
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 DB 1831 TTGGCGTGCCACCTGAGTGTACGCCCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1890
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrIleAspProPheCysValAlaArgCys 600
 DB 1891 GCTGACCAAGTGTGGCTGTGGCCACTATTAAGACCCCTCCCTTCCTGCGTGGCCGCTGC 1950
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu 620
 DB 1951 CCCAGCGGTGAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 DB 2011 GCGCATGCCAGCCTTGCCCCATCACTGCACCCACCTCCTGTGTGGACCTGGATGACAA 2070
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
 DB 2071 GGCTGCCCGCGGACGAGAGCCAGCCCTCTGACGTCCATCGTCTCTGCGGTGGTTGGC 2130
 QY 653 ----- 653
 DB 2131 ATTCTGTGGTCGTGGTCTTGGGGGTGGCTTTTGGGATCCTCATCAAGCGACGACGACAG 2190
 QY 653 ----- 653
 DB 2191 AAGATCCGGAAGTACACGATCGGAGACTGTGTCAGGAAACGGAGCTGGTGGAGCCCGCTG 2250
 QY 653 ----- 653
 DB 2251 ACACCTAGCGGAGCGATGCCCAACCAAGCGCGAGATGCGGATCCTGAAAGACGACGAGCTG 2310
 QY 653 ----- 653
 DB 2311 AGGAAGGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2370
 QY 653 ----- 653
 DB 2371 CCTGATGGGAGAATGTGAAATTCACAGTGGCCATCAAGTGTTCAGGGAAAAACACATCC 2430
 QY 653 ----- 653

DB 2431 CCCAAAGCCAAACAAAGAAATCTTAGCAGGAAGCATACGTGATGGCTGGTGTGGCTCCCCA 2490
 QY 653 ----- 653
 DB 2491 TATGTCTCCCGCCTTCTGGGCATCTGCGCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2550
 QY 653 ----- 653
 DB 2551 ATGCCCTATGCTGCTCTTAGACCATCTCGGGGAAAACCGCGAGGAGCTGGGCTCCAG 2610
 QY 653 ----- 653
 DB 2611 GACCTGTGAATGCTGTATGATGATGCCAAGGGGATGAGCTACCTGGAGGATGCGG 2670
 QY 653 ----- 653
 DB 2671 CTCGTACACAGGACTTGGCCGCTCGGGAACGTGTGTTCAAGAGTCCCAACCATGTCAA 2730
 QY 653 ----- 653
 DB 2731 ATTACAGACTTGGGCTGGCTCGGCTGTGGACATTGACGAGACAGAGTACCATCAGAT 2790
 QY 653 ----- 653
 DB 2791 GGGGCAAGGTGCCCATCAAGTGGATGCGCTGGAGTCCATTCTCCGCGGCGGTTCAAC 2850
 QY 653 ----- 653
 DB 2851 CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2910
 QY 653 ----- 653
 DB 2911 AAACCTTACGATGGATGCCATCCAGCCCGGAGATCCCTGACCTCTCGAAAAGGGGAGCGG 2970
 QY 653 ----- 653
 DB 2971 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAAAATGTTGGATG 3030
 QY 653 ----- 653
 DB 3031 ATTGACTCTGAATGTGCGCCCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCGCAATGGCC 3090
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 DB 3091 AGGGACCCCCAGCGCTTTGTGTCATCCAGAATGAGGACTTGGGCCCCAGCAGTCCCTTG 3150
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 DB 3151 GACAGCACTTCTACCGCTCACTGTGTGGAGACGATGACATGGGGAGCTTGGTGTGATGCT 3210
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
 DB 3211 GAGGAGTATCTGGTACCCACGAGGGCTTCTGTCTCCAGACCTTGGCCCCGGGCGCTGGG 3270
 QY 705 GlyMetValHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
 DB 3271 GGCATGTTCCACACAGCACCACCGCAGCTCATCTACCAGAGTGGGCGTGGGACCTGACA 3330
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 DB 3331 CTAGGGTGGAGCCCTCTGAAGAGAGGAGGCCCCAGGTCTCCACTGGCACCTCCGAAAGG 3390
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
 DB 3391 GCTGGCTCCGATGATTTGATGGTAGCTGGGAATGGGGGAGCGACGAGGGCTGCAAGC 3450
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 DB 3451 CTCCCCACACATGACCCCGCCCTCTACAGCGGTACAGTGAAGAGCCACACAGTACCCTG 3510
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 DB 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal

Db 3511 CCTCTGAGACTGATGGCTACGTTGCCCTGACCTGACGAGCCCGCTGAATATGTG 3570
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
 Db 3571 AACCCAGAGATGTCGGCCCGAGCCCTTGGCCCGAGAGGGCCCTCTGCCCTGTGCTGCC 3630
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
 Db 3631 CGACCTGCTGGTCCCACTCTGGAAGGGCCCAAGACTCTCTCCCGAGGAAGTGGGTC 3690
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
 Db 3691 GTCAAGAGCGTTTTCCTTGGGGTGGCTGGAGAACCCCGAGTACTTGACACCCAG 3750
 QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3751 GAGGAGCTGCCCTCAGCCCACTCTCTCTGCTTCCAGCCAGCTTCGACAACCTC 3810
 QY 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
 Db 3811 TATTACTGGACCAAGACCCAGAGCGGGGGCTCCACCCAGCACCTTCAAAGGGACA 3870
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3871 CTACGGCAGAGAACCAGAGTACCTGGGTCTGGAGCTGCCAGTG 3915

RESULT 7

US-09-167-322-4
 : Sequence 4, Application US/09167322
 : Patent No. 6365151
 : GENERAL INFORMATION:
 : APPLICANT: Allegheny University of the Health
 : Sciences, Halpern, Michael S.
 : England, James M.
 : TITLE OF INVENTION: CANCER VACCINE
 : NUMBER OF SEQUENCES: 14
 : CORRESPONDENCE ADDRESSES:
 : ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
 : STREET: Suite 1800, Two Penn Center Plaza
 : CITY: Philadelphia
 : STATE: PA
 : COUNTRY: USA
 : ZIP: 19102

COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/167,322
 : FILING DATE: 07-Oct-1998
 : CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US97/00582
 : FILING DATE: <Unknown>
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Monaco, Daniel A.
 : REGISTRATION NUMBER: 30,480
 : REFERENCE/DOCKET NUMBER: 7933-33 PC
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (215) 568-8383
 : TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 4:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 4530 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-167-322-4
 Alignment Scores: 0 Length: 4530
 Pred. No.:

Score: 4892.00 Matches: 918
 Percent Similarity: 73.15% Conservative: 0
 Best Local Similarity: 73.15% Mismatches: 1
 Query Match: 96.34% Indels: 336
 DB: 4 Gaps: 1

SEQ6 (1-919) x US-09-167-322-4 (1-4530)

QY	1	MetGluLeuAlaAlaLeuCysArgTTPGlyLeuLeuAlaLeuLeuProProGlyAla	20
Db	151	ATGAGCTGGCGCCCTGTGCCCTGGGGCTCCTCTCGCCCTCTTGGCCCCGGAGCC	210
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	211	GGGAGCACCAAGTGTGCACCGGCACAGACATGAAGTGGGGCTCCTGCCAGTCCGAG	270
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
Db	271	ACCCACTGGACATGCTCGCCACCTTACACAGGCTGCCAGGTGGTGCAGGGAACCTG	330
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	331	GAACTCCTACCTGCTGCCACCAATGCCAGCTGTCTTCTCGAGGATATCCAGAGGTC	390
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	391	CAGGGCTAGCTGCTCATCTCAACCAAGTGAAGTGGCTGCCACTGCCAGAGCTGCGG	450
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	451	ATTGTGCGAGGCAACCACTCTTGGAGCAACTATGCCCTGGCTGTAGACAATGGA	510
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	511	GACCCCTGAACAAATACCACTGTCACAGGGGCTCCCCAGAGGCTGCCGGAGCTG	570
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Db	571	CAGCTTCGAAGCCTCAGAGATCTTGAAGAGGGGTCTTGATCCAGGGGAACCCAC	630
QY	161	LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAla	180
Db	631	CTCTGTACAGGACACGATTTGTGAAGGACATCTCCACAAGAACACACAGCTGGCT	690
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	691	CTCACACTATAGACACCAACCTCTCGGGCTGCCACCTCTTCTCGCATGTGTAAG	750
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	751	GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTGAGAGCTGACGCGCTGCTGT	810
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGluGlnCys	240
Db	811	GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGCTGCCATGACAGAGTG	870
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	871	GCTGCCGGCTGCACGGGGCCCAAGCACTCTGACTGCTGGCCCTGCCCTCCACTTCA	930
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	931	AGTGGCATCTGTGAGCTGCACCTGCCCGACCTGGTCACTACACACACACAGACGTT	990
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	991	TCCATGCCCAATCCCGAGGCGCGTATACATTCGGCGCCAGCTGTGTGCTGTCTGCC	1050
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1051	TACAACTACCTTTCTACGGACGTGGGATCTTGACCCCTCTGTGCCCCCTGCACACAA	1110
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340


```
QY      725  LeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGly  744
         |||||
Db      3331 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCGCCAGGTCTCCACTGGCACCCTCCGAAGGG  3390
QY      745  AlaGlySerAspValPheAspClyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer  764
         |||||
Db      3391 GCTGCTCCGATGATTTGATGCTGACCTGGGAATGGGGGAGCCCAAGGGGCTGCAAAAGC  3450
QY      765  LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu  784
         |||||
Db      3451 CTCGCCACATAGACCCAGCCCTCTACAGCGGTACAGTGGAGACCCACAGTACCCCTG  3510
QY      785  ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal  804
         |||||
Db      3511 CCCTCTGAGACTGATGGCTACGTTGCCCCCTGACCTGCAGCCGCCAGCCTGAATATGTG  3570
QY      805  AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla  824
         |||||
Db      3571 AACCAAGCAGATGTTTCGGCCCCCAGCCCTTTCGCCCCGAGAGGGCCCTCTGCTGCTGCC  3630
QY      825  ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal  844
         |||||
Db      3631 CGACCTGCTGGTGCACCTCTGGAAGGGCCAGACTCTCTCCCGAGGAGAAATGGGTC  3690
QY      845  ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln  864
         |||||
Db      3691 GTCAAAGAGCTTTTTCCTTTGGGGTGCCTGGAGAAACCCGAGTACTTGACACCCAG  3750
QY      865  GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu  884
         |||||
Db      3751 GGAGGAGCTGCCCTCAGCCCCCAGCCCTCTCTGCTCTCAGCCGAGCTTCGACAAACCTC  3810
QY      885  TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr  904
         |||||
Db      3811 TATTACTGGACCAAGGCCACACAGCGGGGGGCTCCACCAGCACCTTCAAAGGGACA  3870
QY      905  ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal  919
         |||||
Db      3871 CCTAGGGCAGAGAACCAGAGTACCTGGGTCTGGAGTGGCCAGTG  3915
```

RESULT 8

US-09-527-487-1

; Sequence 1, Application US/09527487

; Patent No. 6528060

; GENERAL INFORMATION:

; APPLICANT: Nicolette, Charles

; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES

; FILE REFERENCE: 126881309200

; CURRENT APPLICATION NUMBER: US/09/527,487

; CURRENT FILING DATE: 2000-03-16

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 4530

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (151)..(3915)

US-09-527-487-1

Alignment Scores:

Pred. No.:	0	Length:	4530
Score:	4892.00	Matches:	918
Percent Similarity:	73.15%	Conservative:	0
Best Local Similarity:	73.15%	Mismatches:	1
Query Match:	96.34%	Indels:	336
DB:	4	Gaps:	1

SEQ6 (1-919) x US-09-527-487-1 (1-4530)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProProGlyAla 20

```
Db      151  ATGGAGCTGGGGGCTTGTGGCGCTGGGGGCTCTCTCTCGCCCTCTTCCCCCGGAGCC  210
QY      21  AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu  40
         |||||
Db      211  GCGAGCACCAAGTGTGCACCGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCAG  270
QY      41  ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu  60
         |||||
Db      271  ACCCACCTGGACATCTCCGCCACCTCTACACAGGCTGCCAGGTGGTGAGGAAACCTG  330
QY      61  GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal  80
         |||||
Db      331  GAACCTCACCTACTGCCACCAATGCCAGCTCTCTCTCTCTCAGGATATCCAGAGGTG  390
QY      81  GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg  100
         |||||
Db      391  CAGGCTACGTCTCATCGCTCACCAACCAAGTGAGGAGTCCCACTGAGAGGCTGGG  450
QY      101  IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly  120
         |||||
Db      451  ATTTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCGTGTACACAATGA  510
QY      121  AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu  140
         |||||
Db      511  GACCGCTGAACAATATCCACCCCTCTCAGGGGCTCCCAAGAGGCTCGCGGAGCTG  570
QY      141  GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln  160
         |||||
Db      571  CAGCTTCAGAGCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGGGGAACCCACG  630
QY      161  LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla  180
         |||||
Db      631  CTCCTCTACAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACAACAGCTGG  690
QY      181  LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys  200
         |||||
Db      691  CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG  750
QY      201  GlySerArgCysTyrGlySerSerGluAspCysGlnSerLeuThrArgThrValCys  220
         |||||
Db      751  GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTCTCAGAGCTCAGCGGCACTGTCT  810
QY      221  AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys  240
         |||||
Db      811  GCCGTGGCTGTGCCCGCTGCAAGGGGCACTGCCCACTGCTGCTGCTGCTGCTGCTG  870
QY      241  AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis  260
         |||||
Db      871  GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGGCTGCTGCTGCTGCTG  930
QY      261  SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu  280
         |||||
Db      931  AGTGGCATCTGTGAGCTGCATGCCCGCCAGCTTGTACCTTACAACACAGACAGTTT  990
QY      281  SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro  300
         |||||
Db      991  TCCATGCCCAATCCCGAGGGCGGTATACATTTCGGCGCAGCTGTGTGACTGCCCTG  1050
QY      301  TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln  320
         |||||
Db      1051  TACAACCTACCTTCTACGGACCTGGGATCTCCACCTCTGCTGCCCTGTCACCAACCA  1110
QY      321  GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg  340
         |||||
Db      1111  GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCC  1170
QY      341  ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn  360
         |||||
Db      1171  GTGTGCTATGGTCTGGGCATGGAGCATTGCGAGAGGTGAGGGCAGTTACCAGTGCCA  1230
QY      361  IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer  380
         |||||
```

Db	1231	ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTTGGCATTTCTGCCGGAGAGC	1290
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1291	TTTTGATGGGGACCCAGAGCTCCACACACGTCGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1350
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1351	GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGCCCGGACAGCTGCCT	1410
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1411	GACCTCAGCGCTTTCCAGAACCTGCAAGTAATCCGGGACGAATCTGCACAAATGGCGCC	1470
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1471	TACTCGCTGACCTTGCAAGGCTGGGCATCAGCTGGCTGGGCTGGCGCTCACTGAGGGAA	1530
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1531	CTGGGCAGTGAGCTGGCCCTCATCCACATAACACCCACCTCTGCTTCGTGCACACGGTG	1590
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1591	CCCTGGGACCAAGCTCTTTTCGAACCCGACCAAGCTCTGTCTCCACACTGCCAACCGGCCA	1650
Qy	501	GluAspGluCysValGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1651	GAGACGAGTGTGGCGGAGGGCTGGCTGCGCCACCAAGCTGTGCGCCGAGGCACTGC	1710
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1711	TGGGGTCCAGGGCCACCCAGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1770
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1771	GTGGAGAAATGCCAGTACTGTCAGGGGCTCCCAAGGAGTATGTGAATGCCAGGCACTGT	1830
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1831	TTGCCGTGCCACCTGAGTGTACAGCCCAAGTGGCTCAGTGACCTGTTTGGACCGGAG	1890
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1891	GCTGACCAGTGTGGCGCTGTGCCCACTATAAGGACCTCCCTTCTGCGTGGCCGCTGC	1950
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1951	CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTTGGAAAGTTTCCAGATGAGGAG	2010
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2011	GGCGCATGCCAGCTTGCCCATCACTGACCCACCTCCCTGTGTGGACCTGGATGACAAG	2070
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2071	GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCTGCTGCGGTGGTTGGC	2130
Qy	653	-----	653
Db	2131	ATTCTGTGTCGCTGGTCTTTGGGGTGTCTTTGGGATCCTCATCAAGCGACGCGCAGAG	2190
Qy	653	-----	653
Db	2191	AAGATCCGGAAGTACACGATCGGGAGACTGCTGCAGGAAACGAGAGCTGGTGGAGCCGCTG	2250
Qy	653	-----	653
Db	2251	ACACCTAGCGGCGATGCCCAACACGCGCAGATCGGATCCTGAAAGAGACGAGCTG	2310
Qy	653	-----	653
Db	2311	AGGAAGGTGAAGTGTCTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2370
Qy	653	-----	653
Db	2371	CTGTATGGGAGAAATGTGAAAATTTCCAGTGGCCATCAAAGTGTGTGAGGAAAAACACATCC	2430
Qy	653	-----	653
Db	2431	CCCAAGCCAACAAGAAATCTTTAGACGAAGCATACGTGATGGCTGGTGTGGGCTCCCCA	2490
Qy	653	-----	653
Db	2491	TATGTCTCCGCCCTTCTGGGCATCTGCTGCATCCACGCTGCAGCTGGTGACACAGCTT	2550
Qy	653	-----	653
Db	2551	ATGCCCTATGGCTGCCTCTTTAGACCATGTTCGGGAAAAACCGCGGACGCTGGGGTCCAG	2610
Qy	653	-----	653
Db	2611	GACCTGCTGAACGTGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2670
Qy	653	-----	653
Db	2671	CTCGTACACAGGACATTGGCCGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAAA	2730
Qy	653	-----	653
Db	2731	ATTACAGACTTCGGGTGGCTCGGCTGCTGGACATTTGACGAGACAGAGTACCATGCAGAT	2790
Qy	653	-----	653
Db	2791	GGGGCAAGCTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCCGCGCGGTTCACC	2850
Qy	653	-----	653
Db	2851	CACCAGAGTGTGTGTGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2910
Qy	653	-----	653
Db	2911	AAACCTTACGATGGGATCCCGCCAGCCCGGGAGATCCCTGACCTGTCTGGAAGGGGAGCGG	2970
Qy	653	-----	653
Db	2971	CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTGTTGATG	3030
Qy	653	-----	653
Db	3031	ATTGACTCTGAATCTCGGCCAAGATTCCCGGAGTTGGTGTCTGAAATTTCTCCCGCATGGCC	3090
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	3091	AGGACCCCCAGCGCTTTTGGGTCTATCCAGANTGAGGACTTGGGCCAGCCAGTCCCTTG	3150
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3151	GACAGCACCTTCTACCGCTCACCTGCTGGAGACGATGACATGGGGACCTGGTGGATGCT	3210
Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3211	GAGGAGTATCTGGTACCCACAGCAGCGGCTTCTTGTCTCAGACCTTGCCCGCGCGCTGGG	3270
Qy	705	GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr	724
Db	3271	GGCATGGTCCACCACAGCCACCGCAGCTCATCTACAGGAGTGGCGGTGGGGACCTGACA	3330
Qy	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	3331	CTAGGGCTGGAGCCCTCTGAAGAGAGGAGGCCCCAGGTCCTCCACTGGCACCTCCGAAGGG	3390
Qy	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
Db	3391	GCTGGCTCCGATGATTTTGTATGTGACCTGGGAATGGGGCAGCCAAAGGGCTTCAAGAGC	3450

```
QY      765  LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db      3451  CCCCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAGGACCCACAGTACCCCTG 3510
QY      785  ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db      3511  CCCTCTGAGACGATGAGTGGCTGGTCCCTCCCTGACCTGCAGCCCGCCAGCCTGAATATATG 3570
QY      805  AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db      3571  AACCAGCCAGATGTTCCGCCCCAGCCCTTCGCCCCCGAGGCGCCCTCTGCCGTGCTGCC 3630
QY      825  ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db      3631  CGACCTGTGTGGTGCACCTCTGAAAGGGCCAAAGACTCTCTCCCAAGGAAGAAATGGGGTC 3690
QY      845  ValLysAspValPheAlaPheGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db      3691  GTCAAAGAGCTTTTGGCCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCCGAG 3750
QY      865  GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
Db      3751  GGAGGAGTGCCTTCAGCCCAACCTCTCTCTGCTTCAGCCCAAGCCTTCGACAAACCTC 3810
QY      885  TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db      3811  TATTACTGGGACGAGGACCCAGACAGCGGGGGGCTCCACCCAGCACCTTCAAAGGGACA 3870
QY      905  ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db      3871  CCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG 3915
```

RESULT 9

```
US-09-877-177A-11
; Sequence 11, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11
```

```
Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 4 Gaps: 1
```

SE06 (1-919) x US-09-877-177A-11 (1-4530)

```
QY      1  MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db      151  ATGGAGCTGGCGCCTTGTGGCGCTGGGGCTCCCTCTCGCCCTCTTGGCCCCCGGAGCC 210
QY      21  AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      211  GCGAGCACCAAGTGTGACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 270
QY      41  ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db      271  ACCACCTGGACATGCTCGGCCACCTCTACCAAGGGCTGCCAGGTGGTGCAGGGAAACCTG 330
```

```
QY      61  GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspGlnGluVal 80
Db      331  GAACCTACCTACCTGCCACCAATGCCAGCCTGTCTCTTCCCTGCAGGATATCCAGAGGTG 390
QY      81  GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      391  CAGGGCTACGTCGCTCATCGCTCACAAACCAAGTGAGGCAGGTCCCACTCAGAGGCTGGG 450
QY      101  IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      451  ATTGTGCCAGGACCCAGCTCTTTGAGGACAACTATGCCGTGGCGCTGCTAGCAATAGA 510
QY      121  AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      511  GACCGCTGAACAATACCACTCCCTCTCAGAGGGCTCCCCAGGAGGCTCGCGGAGCTG 570
QY      141  GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db      571  CAGCTTCAAGCCTCACAGAGATCTTGAAGAGAGGGTCTTGATCCAGCGGAACCCCGAG 630
QY      161  LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      631  CTCCTCTACAGGACACGATTTTGTGAAGAGGACATCTTCCACAAGAACAACAGCTGGCT 690
QY      181  LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      691  CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG 750
QY      201  GlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      751  GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTCTCAGAGCTGTCAGCGCACTCTCTGT 810
QY      221  AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db      811  GCCGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCCTGCTGCTGCTGCTGCTGCTGCTGCT 870
QY      241  AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      871  GCTGCCGCTGCAGCGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 930
QY      261  SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      931  AGTGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACAGACACGCTTTGAG 990
QY      281  SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      991  TCCATGCCCAATCCCGAGGGCGGTATACATTGGCGCCAGCTGTGTGCTGCTGCTGCTGCT 1050
QY      301  TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      1051  TACAACTACCTTTCTACGGAGCTGGATCCTGCACCCCTGCTGCTGCTGCTGCTGCTGCTG 1110
QY      321  GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      1111  GAGGTGACAGCAGAGAGATGGAACACAGCGGTGTGAGAAAGTGCAGCAAGCCTGTGCCGA 1170
QY      341  ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1171  GTGTGCTATGGTCTGGGCATGAGGACCTTGGCAGAGGTGAGGGCAGTACCAGTGCCTGCT 1230
QY      361  IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      1231  ATCCAGGAGTGTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATCTTCTGCGGAGAGC 1290
QY      381  PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400
Db      1291  TTTGATGGGAGCCCAAGCCTCCAAACACTGCCCGCTCCAGCCAGACAGCTCCCAAGTGT 1350
QY      401  GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db      1351  GAGACTCTGGAGAGATCACAGGTATACCTATACATCTCAGCATGTCGCCGACAGCTGCT 1410
```

QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1411 GACCTCAGCGTCTCCAGAACCTCAAGTAATCCGGGACGAATTCGCACAATGGCGCC 1470
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1471 TACTCGGTGACCCCTGCAAGGGCTGGCATCAGCTGGCTGGGGCTCGGCTCACTGAGGAA 1530
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1531 CTGGGCAGTGACCTGGCCCTCATCCACATACACCCACCTCTGCTCGTCACACAGGTG 1590
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1591 CCCTGGGACCAAGCTCTTTGGGAACCGCACCAAGCTCTGCTCCACACTGGCCAAACCGCCA 1650
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1651 GAGGACGAGTGTGGGCGAGGGCTGGCCCTGCCACCAAGCTGTGCGCCCGAGGCACTGC 1710
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1711 TGGGGTCAGGGCCCCCAGCTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGAGTGC 1770
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTrpValAlaAlaArgHisCys 560
 DB 1771 GTGGAGGAATGCCAGTACTGCAGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT 1830
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 DB 1831 TTGCCGTGCCACCTGAGTGTACGCCCGCAGAAATGGCTCAGTGACCTGTTTTGGACCGGAG 1890
 QY 581 AlaAspGlnCysValAlaCysAlaHisTrpLysAspProPheCysValAlaAlaArgCys 600
 DB 1891 GCTGACCAAGTGTGGCGCTGTGCCCACTATAAGACCCCTCCCTCTCGGTGGCCCGCTGC 1950
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 DB 1951 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 2010
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 DB 2011 GGCCGATGCCAGCCTTGCCCATCACTGCACCCACCTCCTGTGTGGACCTGGATGACAAG 2070
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
 DB 2071 GGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTCCATCGTCTCTGCGGTGGTGGC 2130
 QY 653 ----- 653
 DB 2131 ATCTCTGCTGGTGTGGTCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCGAGCGCAGCAG 2190
 QY 653 ----- 653
 DB 2191 AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTG 2250
 QY 653 ----- 653
 DB 2251 ACACCTAGCGGAGGATGCCCAACCGCGGAGATGCGGATCTCGAAAGAGAGAGGAGCTG 2310
 QY 653 ----- 653
 DB 2311 AGGAAGTGAAGGTGCTGGATCTGCGCTTTTGGCACAGTCTACAGGGCATCTGGATC 2370
 QY 653 ----- 653
 DB 2371 CCTGATGGGAGAAATGTGAATAATCCAGTGGCCATCAAAAGTTGTGAGGGAAAAACACATCC 2430
 QY 653 ----- 653
 DB 2431 CCCAAAGCCAAAGAAATCTTTAGACGAAGCATACGTGATGGTGTGGCTGCCCA 2490
 QY 653 ----- 653

DB 2491 TATCTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2550
 QY 653 ----- 653
 DB 2551 ATGCCCTATGCTGCCTCTTAGACCATGTCCGGGAAAAACCGGACGCTGGGCTCCAG 2610
 QY 653 ----- 653
 DB 2611 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2670
 QY 653 ----- 653
 DB 2671 CTGCTACACAGGAGCTTGGCGCTCGGAACCTGCTGGTCAAGAGTCCCAACCATGTCAAA 2730
 QY 653 ----- 653
 DB 2731 ATTACAGACTTCGGGCTGGCTCGCTGCTGGACATTTGACGAGACAGAGTACCATGCAGAT 2790
 QY 653 ----- 653
 DB 2791 GGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTTCTCCGCCCGGGTTACCC 2850
 QY 653 ----- 653
 DB 2851 CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2910
 QY 653 ----- 653
 DB 2911 AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGGAGCGG 2970
 QY 653 ----- 653
 DB 2971 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAAAATGTTGGATG 3030
 QY 653 ----- 653
 DB 3031 ATTGACTCTGAATCTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCCGCATGSCC 3090
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 DB 3091 AGGGACCCCGAGCGCTTTGGTGCATCCAGAATCAGGACTTTGGGCCCGCAGCAGTCCCTTG 3150
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 DB 3151 GACAGCACCCTTACCCCTCACTGCTGGAGGACCATGACATGGGGGACCTGGTGGATGCT 3210
 QY 685 GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly 704
 DB 3211 GAGGAGTATCTGGTACCCAGCAGCGGCTTCTCTGTCCAGACCTTGCCCGCGCGCTGGG 3270
 QY 705 GlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyAspLeuThr 724
 DB 3271 GGCATGTCTCCACACAGCAGCCAGCTCATCTACAGGAGTGGCGGTGGGAGCTGCACA 3330
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 DB 3331 CTAGGGCTGGAGCCCTCTGAAGAGAGGAGGCCCCAGGTCTCCACTGGCACCTCCGAAGGG 3390
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
 DB 3391 GCTGGCTCCGATGTATTTGATGGTACCTGGGAATGGGGCAGCAAGGGGCTCCAAAGC 3450
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 DB 3451 CTCCCCACATGACCCAGCCCTCTACAGCGGTACAGTGGAGGAGCCACAGTACCCTG 3510
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 DB 3511 CCCTCTGAGACTGATGGCTACGTTGCCCCCTGACCTGCAGCCCCCAGCCCTGTAATATGTG 3570
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824

Db 3571 AACACGCGCAGATGTTGCGGCCCCAGCCCCCTTCGCCCGAGAGGGCCCTCTGCGCTGCTGCC 3630
Qy 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3631 CGACCTGCTGGTCCACTTGGAAAGGGCCAGACTCTCTCCCGAGGAGAAATGGGGTC 3690
Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3691 GTCAAGAGCTTTTTCCTTGGGGTGGCGGTGAGAACCCCGAGTACTTGACACCCAG 3750
Qy 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3751 GGAGGAGTGGCCCTCAGCCCCACCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3810
Qy 885 TyrTyrTriAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3811 TATTACTGGACGAGCCACCAGAGCGGGGGCTCCACCCAGACCTTCAAGGGGACA 3870
Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3871 CCTACGGCAGAGAACCAGTAGTACCTGGGTCTGGAGTGGCCAGTG 3915

RESULT 10

US-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPlastic DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-229-515A-14

Alignment Scores:
Pred. No.: 9,44e-287 Length: 3955
Score: 4134.00 Matches: 781
Percent Similarity: 65.47% Conservative: 42
Best Local Similarity: 62.13% Mismatches: 96
Query Match: 81.41% Indels: 338
DB: 1 Gaps: 3

SEQ6 (1-919) x US-08-229-515A-14 (1-3955)

Qy 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 26 ATGGAGCTGGCGGCTGGTGGCGCTGGGGTCTCTCTCGCCCTCTCTCGCCCGGGAATC 85
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 86 GCGGGCACCAAGTGTGTACCGGCACAGACATGAAGTTGGGCTCTCTGCCACTCTCTGAG 145
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu 60
Db 146 ACCCACTGGACATGTCTCGCCACCTGTACCAAGGGCTGTCTAGTGTAGGAGCACTTG 205
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 206 GAGCTTACCTACGTGCTGCAATGCCAGCTCTCATCTCTCTGAGGACATCCAGGAAGTT 265
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 266 CAGGGTTACATGCTCATGCTCAACACCAAGGTGAAGCGGTCCCACTGCAAGAGCTGCGC 325
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 326 ATCGTGAGAGGACCCAGCTCTTTGAGGACAAGTATGCCCTGCTGTAGACAACCGA 385
Qy 121 AspProLeuAsnAsnThrThrProValThr---GlyAlaSerProGlyGlyLeuArgGlu 139
Db 386 GATCCTCAGGACAATGTCGCGGCTCCACCCAGGAGAACCCAGAGGGGCTGCGGGAG 445
Qy 140 LeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnPro 159
Db 446 CTGCAGCTTCGAAGTCTCAGAGATCTCTGAAGGGAGAGTTTGTATCGGTGGGAACCT 505
Qy 160 GlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeu 179
Db 506 CAGCTCTGTACCAGGACATGTTTGTGGAAGGAGCTCTTCGCAAGATAACCAACTG 565
Qy 180 AlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCys 199
Db 566 GCTCTCTCGATATAGACACCAATCTTCCGGGCTGTCCACCTCTGCTGCCCCGCTGC 625
Qy 200 LysGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrVal 219
Db 626 AAAGACAATCACTGTTGGGTGAGTCCGGAAGACTGTGAGATCTTGACTGGCCACATC 685
Qy 220 CysAlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGln 239
Db 686 TGTACCAGTGGTTGTCGCCGTCGCAAGGCGGCTGCCCACTGCTGCTCCATGAGCAG 745
Qy 240 CysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsn 259
Db 746 TGTGCCGAGGCTGCACGGGCCCCAAGCATTTCTGACTGCGCTGCGCTGCCCTCAAT 805
Qy 260 HisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPhe 279
Db 806 CATAGTGGTATCTGTGAGTGTGCTGCCAGCCCTGTACCTTACCAACACAGACACTTT 865
Qy 280 GluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCys 299
Db 866 GAGTCCATGCACAACCTGAGGGTCTGCTACACCTTTGGTGGCAGCTGCGTGACCCTGC 925
Qy 300 ProTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsn 319
Db 926 CCTACAACCTACCTGTCTAGGAAGTGGATCTCTGCACTCTGCTGCTGCGCGGAATAAC 985
Qy 320 GlnGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAla 339
Db 986 CAAGAGGTACAGCTGAGGACGGAACACAGCGTTGTGAGAAATGCAAGAGCCCTGGCT 1045
Qy 340 ArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAla 359
Db 1046 CGAGTGTCTATGCTGTGGCATGAGCACCTTTCGAGGGGCGAGGCCCATCCAGTGCAC 1105
Qy 360 AsnIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGlu 379

QY 763 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 782
 Db 3326 CAGAGGCTCTCCACATGACCTCAGCCCTCTACAGGGGTACAGGAGGACCCACATTA 3385
 QY 783 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlu 802
 Db 3386 CCTCTGCCCGCCGAGACTGATGGCTATGTTGGTCCCTGGCTGCGAGCCCGAG 3445
 QY 803 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 822
 Db 3446 TATGTGAACCAATCAGAGGTTTCAGCCTCAGCCTCTTAAACCCAGAGGGTCTCTGCT 3505
 QY 823 AlaAlaArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsn 842
 Db 3506 CCTGTCCGCGCTCTGGTCTACTCTAGAAACCCAGACTCTCTCTCTCTGGGAGAT 3565
 QY 843 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 862
 Db 3566 GGGGTGTCAAGAGGTTTTTTCCTTCGGGGGTGCTGTGGAGAACCTGAATACTAGTA 3625
 QY 863 ProGlnGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 882
 Db 3626 CCGAGAGAGGCACTGCTCTCCGCCCCACCTCTCTCTGCTTCCAGCCAGCCTTTGAC 3685
 QY 883 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 902
 Db 3686 AACCTCTATTACTGGGACCAAGACTCATCGGAGAGGCGGCTCCACCAAGTAACCTTGAA 3745
 QY 903 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3746 GGGACCCCACTGCAGAACCCCTGAGTACCTAGGCTAGGCTGGATGTACCTGTA 3796

RESULT 11

US-08-645-865-14
 ; Sequence 14, Application US/08645865
 ; Patent No. 5654406
 ; GENERAL INFORMATION:
 ; APPLICANT: RAZIUDIN
 ; APPLICANT: SARKAR, FAZLUL H
 ; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
 ; TITLE OF INVENTION: NEOPLASTIC DISEASE
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NEEDLE & ROSENBERG PC
 ; STREET: 127 Peachtree Street, Suite 1200
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: usa
 ; ZIP: 30303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/645,865
 ; FILING DATE: 14 MAY 1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PERRYMAN, DAVID G
 ; REGISTRATION NUMBER: 33,438
 ; REFERENCE/DOCKET NUMBER: 1414.608
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 404-688-0770
 ; TELEFAX: 404-688-9880
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3955 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-645-865-14

Alignment Scores:
 Pred. No.: 9,44e-287 Length: 3955
 Score: 4134.00 Matches: 781
 Percent Similarity: 65.4% Conservatives: 42
 Best Local Similarity: 62.1% Mismatches: 96
 Query Match: 81.4% Indels: 338
 DB: 1 Gaps: 3
 SEQ6 (1-919) x US-08-645-865-14 (1-3955)
 QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProGlyAla 20
 Db 26 ATGGAGCTGGCGGCTGGTGGCGCTGGGGTTCCTCTCGCCTCTCTCCGCCCGGAATC 85
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProGlu 40
 Db 86 GCGGCGACCAAGTGTACCGGCACAGACATGAAGTTCGGCTCTCCCTGCCAGTCTCTAG 145
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 146 ACCCACCCTGGACATGCTCGGCCACCTGTACACAGGCTGTACAGTGTGTCAGGGCACTTG 205
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 206 GAGCTTACCTAGCTGCTGCCAATGCCAGCCTCTCATTCCTGCAGGACATCCAGGAAGTT 265
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 266 CAGGGTTACATCTCATCGCTCACAAACAGGTGAAGCGCGTCCCACTGCAAGGTCGCGC 325
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 326 ATCGTGAGAGGACCCAGCTCTTTGAGGACAGTATGCCCTGGCTGTCTACACACCGA 385
 QY 121 AspProLeuAsnAsnThrThrProValThr----GlyAlaSerProGlyGlyLeuArgGlu 139
 Db 386 GATCCTCAGGACAATGTGCGCGCTCCACCCAGGACAGAACCCAGAGGGCTGCGGGAG 445
 QY 140 LeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnPro 159
 Db 446 CTGCAGCTTCGAAGTCTCACAGATCTCTGAAGGAGGAGTTCCTCCGCAAGAAATAACCAACCT 505
 QY 160 GlnLeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeu 179
 Db 506 CAGCTCTGCTACACGAGACATGCTTTTGTGGAAGAGCTTCCTCCGCAAGAAATAACCAACCT 565
 QY 180 AlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCys 199
 Db 566 GCTCCTGTGATATAGACACCAATGTTCCCGGGCTGTCTCCACCTTGTGCCCGCGCTGC 625
 QY 200 LysGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrVal 219
 Db 626 AAAGACAATCACTGTTGGGGTGAGAGTCGGGAAGACTGTTCAGATCTTTGACTGGCACCATC 685
 QY 220 CysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGln 239
 Db 686 TGTACCACTGTTGTCGGGTGCAAGGCGCGGCTGCCACTGTCTGCTGCCATGAGCAG 745
 QY 240 CysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsn 259
 Db 746 TGTGCCGAGGCTGCACGGGCCCAAGCATCTGACTGCTGCTGCTGCTGCCCTCAAT 805
 QY 260 HisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPhe 279
 Db 806 CATAGTGTATCTGTGAGCTGCACCTGCCAGCCCTCGTACCTACACACACACACACCTTT 865
 QY 280 GluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCys 299
 Db 866 GAGTCCATGCACACCCCTGAGGGTCTGCTACACCTTTGGTGGCCGCTGCTGCTGCTGCTG 925
 QY 300 ProTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsn 319
 Db 925

```
Db 926 CCCTACAACTACCTGTCTACCGAAGTGGGATCCCTGCACCTCTGGTGTGTCCTCCCGGAATAAC 985
Qy 320 GlnGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAla 339
Db 986 CAAGAGGTTCACAGCTGAGGACGGAACACAGCGTTGTGAGAAATGCAGCAACCCCTGTGCT 1045
Qy 340 ArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAla 359
Db 1046 CGAGTGTGTATGCTGTGGGATGAGCACCTTCAGGGGCGAGGGCCATCACCAGTGAC 1105
Qy 360 AsnIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGlu 379
Db 1106 AATGTCAGGAGTTTGATGGCTGCAAGAAGATCTTTGGAGCCCTGGGCAATTTTGCCTGGAG 1165
Qy 380 SerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnVal 399
Db 1166 AGCTTTGATGGGACCCCTCCTCCGGCATTCCTCCGCTGAGGCCCTGAGCAGCTCCNAGTG 1225
Qy 400 PheGluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeu 419
Db 1226 TTCGAAACCCCTGGAGGAGATCACAGGTACCTGTACATCTCAGCATGCCAGACAGTCTC 1285
Qy 420 ProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGly 439
Db 1286 CGTGACCTCACTGTCTTCCAGAACCTTCGAATCATTCGGGGACGGATCTCCACGATGGC 1345
Qy 440 AlaTyrSerLeuThrLeuGlnGlyIleSerTrpLeuGlyLeuArgSerLeuArg 459
Db 1346 CGGTACTATTGACATGCAAGGCTGGGGATCCACTCGCTGGGGCTGCGCTCACTGCGG 1405
Qy 460 GluLeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThr 479
Db 1406 GAGCTGGGCAGTGATGGCTCTGATTCAACGCAACGCCCATCTCTGCTTTGTACACACT 1465
Qy 480 ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg 499
Db 1466 GTACCTTGGGACCAAGCTCTTCCGGAACCCACATCAGGCCCTGCTCCACAGTGGGAACCGG 1525
Qy 500 ProGluAspGlu---CysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGly 518
Db 1526 CCGGAAGAGGACTTGTGCGTCTCGAGCGCTTGGTCTGTACTCACTGTGTGTCGCCACGGG 1585
Qy 519 HisCysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGln 538
Db 1586 CACTGCTGGGGCCAGGGCCACCACGAGTGTCAACTGCAGTCACTTCTTCGGGGCCAG 1645
Qy 539 GluCysValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArg 558
Db 1646 GAGTGTGTGGAGGAGTCCCGAGTATGGAAGGGGCTCCCGGGAGTATGTGAGTGACAAG 1705
Qy 559 HisCysLeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGly 578
Db 1706 CGCTGTCTGCGGTGTACCCCGAGTGTGAGCTTCAAAACAGCTCAGAGACCTGCTTTTGA 1765
Qy 579 ProGluAlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAla 598
Db 1766 TCGGAGGCTGATCAGTGTGACGCTGCGCCCACTACAAGGACTCGCTCTCTGTGTGCT 1825
Qy 599 ArgCysProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAsp 618
Db 1826 CGCTGCCCCAGTGTGTGAACCGGACCTCTCTACATGCCCCATCTGGAAGTACCCGGAT 1885
Qy 619 GluGluAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAsp 638
Db 1886 GAGAGGGCATATCCAGCCGTGCCCCATCACTGACACCACTCCTCTGTGTGGATCTGGAT 1945
Qy 639 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThr----- 652
Db 1946 GAACGAGGCTGCCCCAGCAGACAGAGCCAGCCCGGTCGATTCATTCATTGCAACTGTA 2005
Qy 652 ----- 652
Db 2006 GAGGGCGTCTGTCTTCTCTGATCTTAGTGGTGGTCTGTGGAATCCTTAATCAACGAAGG 2065
```

```
Qy 652 ----- 652
Db 2066 AGACAGAAGATCCCGAAGTATACGATCGTAGGCTGCTGCAGGAAACTGACTTAGTGGAG 2125
Qy 652 ----- 652
Db 2126 CCGCTGACGCCACGCGGAGCAATGCCAACCAAGGCTCAGATGCGGATCCTTAAAGAGACG 2185
Qy 652 ----- 652
Db 2186 GAGCTAAGGAAGGTGAAGTCTTGGATCAGGAGCTTTTGGCACTGTCTACAAGGGCATC 2245
Qy 652 ----- 652
Db 2246 TGGATCCAGATGGGAGAAATGTGAAAATCCCCGTGGCTATCAAGGTGTTGAGAGAAAAC 2305
Qy 652 ----- 652
Db 2306 ACATCTCCTAAAGCCAAACAAAGAAATTCATAGATGAAGCGTATGTATGGTGGTGGGT 2365
Qy 652 ----- 652
Db 2366 TCTCCGATATGTCTCCCGCTCCTGGGCATCTGCTGACATCCACAGTACAGCTGCTGAC 2425
Qy 652 ----- 652
Db 2426 CAGCTTATGCCCTACGGCTGCCCTTCTGGACCATGTCCGAGAACACCGAGGTGCGCTAGGC 2485
Qy 652 ----- 652
Db 2486 TCCCAGGACCTGCTCAACTGGTGTTCAGATTGCCAAGGGGATGAGCTACCTGGAGGAC 2545
Qy 652 ----- 652
Db 2546 GTGCGGCTGTACACAGGGACCTGGCTGCCCGGAATGTGCTAGTCAAGAGTCCCAACCCAC 2605
Qy 652 ----- 652
Db 2606 GTCAAGATTACAGATTTCCGGCTGGCTCGGCTGGACATTTGATGACAGAGATACCCT 2665
Qy 652 ----- 652
Db 2666 GCAGATGGGGCAAGGTGCCCATCAAATGGATGGCATTTGGAATCTATTCTCAGACGCGG 2725
Qy 652 ----- 652
Db 2726 TTCACCATCAGAGTGTGTGTGGAGCTATGGAGTGTGTGGGAGCTGATGACTTTT 2785
Qy 652 ----- 652
Db 2786 GGGGCCAAACCTTACGATGGAAATCCAGCCCGGAGATCCCTGATTTTCTCGGAGAGGGA 2845
Qy 652 ----- 652
Db 2846 GAAGCGCTACCTCAGCCCTCCAATCTGCACCATTTGATGTCTACATGATTATGCTCAAAATGT 2905
Qy 652 ----- 652
Db 2906 TGGATGATTGACTCTGTAATGTGCGCCGAGATTCGGGGAGTTGGTGTCTCAGAAATTTTCAGT 2965
Qy 653 -----SerGlnAsnGluAspLeuGlyProAlaSer 662
Db 2966 ATGGCGAGGAGACCCCGAGCGTTTGTGTGTCATCAGAAACGAGGACTTTGGGCCCATCAGC 3025
Qy 663 ProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuVal 682
Db 3026 CCCATGGACAGTACCTTCTACCGTTCACTGCTGGAAGATGATGACATGGGTGACCTGGTA 3085
Qy 683 AspAlaGluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGly 702
Db 3086 GAGCCTGAAGAGTATCTGGTGGCCCGCAGCAGGGATTTCTTCCCGCGGAGCCCTACCCAGGC 3145
```

```
QY 703 AlaGlyGlyMetValHisArgHisArgSerSerSerThrArgSerGlyGlyGlyAsp 722
Db 3146 ACTGGGAGCAGACCCATAGAAGCCACCAGCTCTCCACCAGGAGTGGAGTGGTGAG 3205
QY 723 LeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer 742
Db 3206 CTGACACTGGGCTGGAGCCCTCGGAAGAAGGGCCCCCAGATCTCCACTGGCTCCCTGG 3265
QY 743 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeu 762
Db 3266 GAAGGGCTGGCTCCGATGTTGATGTGACCTGGCAATGGGGTACCAAGAGGCTG 3325
QY 763 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 782
Db 3326 CAGAGCTCTCTCCACACTGACCTCAGCCCTCTACAGCGGTACAGCGAGGACCCACATTA 3385
QY 783 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlu 802
Db 3386 CCTCTGCCGCCGAGACTGATGGCTATGTTGCTCCCTGGCCTGCAGCCCCCAGCCCGAG 3445
QY 803 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 822
Db 3446 TATGTGAACCATCAGAGGTTGAGCTCAGCTCAGCTCTTTAAACCCAGAGGGTCTCTGCT 3505
QY 823 AlaAlaArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsn 842
Db 3506 CCTGTCCGGCTGCTGGTGTACTCTAGAAAGACCCAGACTCTCTCTCTGGGAAGAT 3565
QY 843 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 862
Db 3566 GGGGTTGTCAAGACGCTTTTGGCTTCGGGGGTGCTGTGGAGAACCCTGAATACTTAGTA 3625
QY 863 ProGlnGlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAsp 882
Db 3626 CCGAGAGAGGACTGCTCCCTCCGCCGCCACCTCTCTCTGCTTCCAGCCAGCCCTTGAC 3685
QY 883 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 902
Db 3686 AACCTCTATTACTGGGACCAAGAACTCATCGGAGCAGGGGCTCCACCAAGTAACCTTGAA 3745
QY 903 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3746 GGGACCCCTCACTCAGAGAACCCTGAGTACCTAGGCTGGATGATACCTGTA 3796

RESULT 12
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
```

```
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; US-09-146-283-3

Alignment Scores:
Pred. No.: 4,4e-251 Length: 2385
Score: 3632.00 Matches: 659
Percent Similarity: 98.65% Conservative: 0
Best Local Similarity: 98.65% Mismatches: 5
Query Match: 71.52% Indels: 4
DB: 2 Gaps: 2

SEQ6 (1-919) x US-09-146-283-3 (1-2385)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 11 ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCTCTCGCTCTCTTGGCCCCCGAGGC 70
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 71 GCGAGCACCAAGTGTGCACCGCACAGACATGAAGCTCGGCTCCCTGCCAGTCCCGAG 130
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 131 ACCCACCTGGACATGTCGCCACCTCTACACAGGCTGCAGGTGTGTCAGGAAACCTG 190
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAluVal 80
Db 191 GAACCTACCTACCTGCCACCAATGCCAGCTGTCTTCTTCAGGATATCCAGGAGGTG 250
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 251 CAGGGCTAGCTGCTCATCGCTCACACCAAGTGAAGCAGCTCCACTGCAGAGGCTGCGG 310
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 311 ATTTGCGAGGCACCCAGCTCTTTGAGGACAACCTATGCCCTGCCCTGTAGACAATGGA 370
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 371 GACCCGCTGAACAATACCACTGCTGTACAGAGGCTCTCCAGGAGGCTTGGGGAGCTG 430
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 431 CAGCTTCGAAGCTCACAGAGATCTTGAAGAGAGGGTCTTGTATCCAGCGGAAACCCAG 490
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db 491 CTCTGCTACACAGACACGATTTTGTGAAGGACATCTTCCAAAGAACCAACAGCTGGCT 550
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 551 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAG 610
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 611 GGCTCCCGCTGCTGGGAGAGAGTCTGAGGATTGTGAGAGCTGTGACGCGCCTGCTGT 670
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGluGlnCys 240
```

```
Db 671 GCGGGTGGCTGTGCCCCCTGCAAGGGGCCACTGCCACTGCTGCCATGAGCAGTGT 730
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 731 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCGCTGGCCCTCCACTTCAAGCAC 790
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 791 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTTACAACACAGACAGCTTTGAG 850
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 851 TCATGCCCCAATCCGAGGGCCGGTATACATTGCGGCCAGCTGTGTGACTGCTGTGCC 910
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 911 TACAACCTACCTTTCTACGACGCTGGGATCTGCACCCCTGCTGCCCTGCACACACCAA 970
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 971 GAGGTGACACAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGAGCCCTGTGCCGA 1030
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1031 GTGTGTATGCTGTGGGCATGGAGCACCTTGCAGAGGTGAGGCAGTTACCACTGGCAAT 1090
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1091 ATCCAGAGTTTCTGGCTGCAGAAGATCTTTGGGAGCTGTCATTTCTGCCGGAGAGC 1150
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1151 TTTGATGGGGACCCAGCCTCCACACACTGCCCCCTCCAGCCAGCAGCTCCAAGTGT 1210
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro 420
Db 1211 GAGACTCTGGAAGAGATCAAGGTTACTATACATCTCAGCATGGCGGACAGCCTGCCT 1270
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1271 GACCTCAGCGCTTCCAGAACCTGCAAGTAAATCCGGGGAGCAATCTGCACATGGCGCC 1330
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1331 TACTCGCTGACCTGCAAGGGCTGGCATCAGCTGGCTGGGGCTGCCTCAGTACAGGGAA 1390
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1391 CTGGGCGAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTGCACACGGTG 1450
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1451 CCTGGGACCAAGCTCTTTGGAAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1510
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1511 GAGGACGAGTGTGGGCGAGGGCCTGGCTGCCACCAAGCTGTGCGCCGAGGGGCACTGC 1570
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1571 TGGGGTCCAGGGGCCACCCAGTGTCAACTGCAGGCAGTTCTCTGGGGCCAGGAGTGC 1630
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1631 GTGGAGAAATGCCAGTACTGCAAGGGCTCCCGAGGGAGTATGTGAATGCCAGGCACTGT 1690
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1691 TTCCCGTGCACCCCTGAGTGTACAGCCCGAGAAATGGCTCACTGACCTGTTTGGACCGGAG 1750
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
|||||
```

```
Db 1751 GCTGACCAAGTGTGGCCTGTGCCCACTATAGGACCCCTCCCTTCTGCTGGCCGCTGC 1810
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1811 CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAAGTTTCCAGATGAGGAG 1870
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1871 GCGCATGCCAGCCTTGCCCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1930
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGlyPro 660
Db 1931 GGCTGCCCCCGAGCAGAGAGCCGCTCTGACGTCCCTCGAG-----GCACCC 1981
QY 661 Ala---SerProLeuAspSerThr 667
Db 1982 GCGCGTCCGCCCAAGCCCAAGCACA 2005
```

```
RESULT 13
US-08-579-823A-3
; Sequence 3, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-08-579-823A-3
```

```
Alignment Scores:
Pred. No.: 4.4e-251 Length: 2385
Score: 3632.00 Matches: 659
Percent Similarity: 98.65% Conservative: 0
Best Local Similarity: 98.65% Mismatches: 5
Query Match: 71.52% Indels: 4
DB: 3 Gaps: 2
```



```

;
; TITLE OF INVENTION: Immunostimulatory Compositions
;
; NUMBER OF SEQUENCES: 10
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Dehlinger & Associates
;
; STREET: 350 Cambridge Ave. Suite 250
;
; CITY: Palo Alto
;
; STATE: CA
;
; COUNTRY: USA
;
; ZIP: 94306
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/09/344,195
;
; FILING DATE: 24-Jun-1999
;
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER: US/09/146,283
;
; FILING DATE: 03-SEPT-1998
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Judge, Linda R.
;
; REGISTRATION NUMBER: 42,702
;
; REFERENCE/DOCKET NUMBER: 7636-0010.21
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: 650-324-0880
;
; TELEFAX: 650-324-0960
;
; INFORMATION FOR SEQ ID NO: 3:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 2385 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: double
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: DNA (genomic)
;
; HYPOTHETICAL: NO
;
; ANTI-SENSE: NO
;
; ORGANISM: homo sapiens
;
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
;
; US-09-344-195-3

```

```

Alignment Scores:
Pred. No.: 4.4e-251 Length: 2385
Score: 3632.00 Matches: 659
Percent Similarity: 98.65% Conservative: 0
Best Local Similarity: 98.65% Mismatches: 5
Query Match: 71.52% Indels: 4
DB: 3 Gaps: 2

SEQ6 (1-919) x US-09-344-195-3 (1-2385)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 11 ATGAGCTGGGGCTGTGCGGCTGGGGCTCCCTCGGCCCTTTGCCCGGAGCC 70
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetIysLeuArgLeuProAlaSerProGlu 40
Db 71 GCGAGCACCCCAAGTGTGCACCGGCACAGACATGAAGCTGGGGCTCCCTGCCAGTCCCGAG 130
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 131 ACCACCTGGAGATGCTCCGCCACCTCTACACAGGGCTGCCAGGTGGTGACGGGAACCTG 190
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 191 GAACCTACCTACCTGCCACCAATGCCAGCGCTGCTTCCCTGCGAGGATATCCAGGAGGTG 250
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 251 CAGGGCTACGTGCTCATCGCTCACAAACCAAGTAGGAGGTGCCCACTGCCAGGCTGCGG 310

```

```

Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 311 ATTTGGGAGGACCCAGCTCTTTGAGGACAATATGCTCCCTGGCGTCTAGACAATGGA 370
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 371 GACCCGCTGAACAATACCAACCCCTGTACAGGGGCTCCCGAGGAGGCTGCGGAGGTG 430
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 431 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGTATCCAGCGGAACCCCGAG 490
Qy 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 491 CTCTGCTACCAAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACCAACAGCTG 550
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 551 CTCACACTGATAGACACCAACCCGCTCTCGGGGCTGCCACCCCTGTTCTCCGATGTGAAG 610
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 611 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGAGATTGTACAGCCTGACCGGCACTGCTGT 670
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 671 GCCGGTGGCTGTGCCCGCTGCAAGGGGCACTGCCCACTGACTGCTGTCATGAGCAGTGT 730
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 731 GCTCCGCTGCACGGGCCCCAAGCACTCTGACTGCTGCTGGCTGCCCTCCATCAACCCAC 790
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 791 AGTGGCATCTGTGAGCTGCACTGCCACGCCCTGGTCACTAACACACAGACACAGTTGAG 850
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 851 TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTCCC 910
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 911 TACAACCTACCTTTCTAGCGAGCTGGATCTCTGCACTGCTGCTGCTGCCCTGCACACCAA 970
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGlnLysCysSerLysProCysAlaArg 340
Db 971 GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1030
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValAlaArgAlaValThrSerAlaAsn 360
Db 1031 GTGTGCTATGCTTGGCATGGACACTTGGCAGAGGTGAGGGCAGTTACCAAGTCCCAAT 1090
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1091 ATCCAGGAGTTTGTGCTGCAGAGAATCTTTTGGAGCCTTGGCATTTCTGCCGAGAGC 1150
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1151 TTTGATGGGACCCAGCCTCCACACTGCCCGCTGCCAGCCAGCAGCAGCTCCAAGTGT 1210
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpAspSerLeuPro 420
Db 1211 GAGACTTGGGAAGAGATCACAGGTACCTATACATCTCAGCATGGCCGAGCAGCTGCCT 1270
Qy 421 AspLeuSerValPhePheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1271 GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGACGAAATCTGCACAAATGGG 1330
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1331 TACTCGTGCCTTGCAGGGCTGGCATCAGCTGGTGGGCTGCGCTCCTCCTCAGGAGGAA 1390
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480

```

```

1391 CTGGCAGTGGAGTGGCCCTCATCCACCAATAACCCACCTCTGCTTGTGCACACGGTG 1450
481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
1451 CCTGGGACAGCTCTTTTGGAAACCGCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1510
501 GluAspGlnCysValGlyGluGlnAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
1511 GAGGACAGTGTGGGCGAGGGCTGGCTGCCACCACTGTGCCCGCCGAGGGCACTGC 1570
521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
1571 TGGGTGCCAGGCGCCACCCAGTGTCAACTGCAGCCAGTCTCTTCGGGGCCAGGAGTGC 1630
541 ValGluGlnCysArgValLeuGlnGlyLeuProArgGluTyrValAlaAsnAlaArgHisCys 560
1631 GTGGAGGAATGGCGAGTACTGAGAGGGCTCCCGAGGGAGTATGTGAATGCCAGGCACTGT 1690
561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
1691 TTGCGGTGCCACCTCAGTGTGAGCCCGCCAGATGGCTCAGTGACCTGTTTGGACGGAG 1750
581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
1751 GCTGACCACTGTGGCTGTGCCACTATAAGGACCTCCCTTCTGCTGGCGCCGCTGC 1810
601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGlu 620
1811 CCAGCGGTGTGAACCTGACCTCTCTATATGCCCATCTCGAAGTTTCCAGATGAGGAG 1870
621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
1871 GCGCATGCCAGCTTGGCCCATCACTGCACCCACTCTCTGTGTGACCTGGATGACAAG 1930
641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGlyPro 660
1931 GGCTGCCCGCGCAGCAGAGAGCAGCCCTCTGACCTGCTCCCTGAG-----GCACCC 1981
661 Ala---SerProLeuAspSerThr 667
1982 GCGCGTCCGCCAGCCCGCCAGCACA 2005

```

RESULT 15

```

US-08-422-108-2
: Sequence 2, Application US/08422108
: Patent No. 6015567
: GENERAL INFORMATION:
: APPLICANT: Hudziak, Robert M.
: APPLICANT: Shepard, H. Michael
: APPLICANT: Ullrich, Axel
: TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WinPatIn (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/422,108
: FILING DATE: 14-Apr-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/355460
: FILING DATE: 13-DEC-1994
: PRIOR APPLICATION DATA:

```

```

: APPLICATION NUMBER: 08/048346
: FILING DATE: 15-APR-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/354319
: FILING DATE: 19-MAY-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: Lee, Wendy M
: REGISTRATION NUMBER: 00,000
: REFERENCE/DOCKET NUMBER: 554C2D2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/223-1994
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1872 nucleotides
: TYPE: Nucleic Acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: US-08-422-108-2

Alignment Scores:
Pred. No.: 1,08e-239 Length: 1872
Score: 3471.00 Matches: 623
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 68.35% Indels: 0
DB: 3 Gaps: 0

SEQ6 (1-919) x US-08-422-108-2 (1-1872)

QY 22 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 41
DB 1 AGACCCAGUGUGACCGGCACAGACAUAGAGGUGCGGCUCCUCCAGUCCCGAGAGACC 60
QY 42 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 61
DB 61 CACUGGACAUUGCUGCCGCCACCUACACAGGCGUGGUGGUGGAGGAGGAGGAGGAG 120
QY 62 LeuThrTyrProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln 81
DB 121 CUCACCUACUGCCCGCCACCAUGCCCGUCCUCCUCCUCCUCCUCCUCCUCCUCCUCC 180
QY 82 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 101
DB 181 GGUACGUGUCUACUGCUCACACCAAGUGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY 102 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 121
DB 241 GUGCGAGGACCCAGCUCUUGAGGACAAUUGCCCGUGGUGGUGGUGGUGGUGGUGGAG 300
QY 122 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlnLeuGln 141
DB 301 CCGUGAACAUAUACCCACCCUGUCACAGGGGCUCCCGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 142 LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 161
DB 361 CUUCGAAGCCUCACAGAGAUUUGAAAGAGGGGUGUUGAUCCAGCGGAAACCCCGAGCUC 420
QY 162 CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAlaLeu 181
DB 421 UGUACACGAGACAGAUUUGGAGGAGCAUCCUCCACAAAGAACACAGGUGGUGCUC 480
QY 182 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 201
DB 481 ACACUGAUAGACACCAACCCUGUCUGGGGCGGCCACCCUCCUCCUCCUCCUCCUCCUCC 540
QY 202 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 221
DB 541 UCCCGCUGCGGGGAGAGAUUCUGAGGAGUUGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 222 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAla 241

```


Db 601 GGUGGUCUGCCCGUGCAAGGGCCACUGCCACUGACUGCCCAUGAGCAGUGUGCU 660
Qy 242 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 261
Db 661 GCCGGUCGACGGGCCCAAGCACUCUGACUGCCUGGCGUCUCCACUUAACACACAGU 720
Qy 262 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 281
Db 721 GGCAUCUGUGAGCUGCAGUGCCAGCCCGUGGUCACCUACAAACACAGACACGUUUGAGUCC 780
Qy 282 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 301
Db 781 AUGCCCAUCCCGAGGGCCGGUAUACAUCCGGCCAGCUGUGUGACUGCCUCCUAC 840
Qy 302 AsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnGlu 321
Db 841 AACUACCUUUCUACGGAGUGGAUCCUGCAGCCUGCUGCCUCCUGCACCAACAGAG 900
Qy 322 ValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArgVal 341
Db 901 GUGACAGCAGAGGAUGGAACACAGCGGUGUGAGAGUGCAGCAAGCCUGUGCCGAGUG 960
Qy 342 CysTyrGlyLeuGlyMetGluHisLeuArgGluValAlaValThrSerAlaAsnIle 361
Db 961 UGCUAUGGUCUGGCGAGGAGCACUUGCGAGGAGGAGGAGGAGUUAACAGUGCCAAUADC 1020
Qy 362 GlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSerPhe 381
Db 1021 CAGGAGUUGGUCGCGCAAGAACAUUUUGGAGCCUGGCAUUUCGCCGAGAGCUUU 1080
Qy 382 AspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPheGlu 401
Db 1081 GAUGGGACCCAGCCUCCACACUGCCCGCUGCCAGCAGCAGCAGCAGCAGCAGCAGCAG 1140
Qy 402 ThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuProAsp 421
Db 1141 ACUCUGGAAGAGAUACAGGUUACCUAUACAUCCAGCAGCAGCAGCAGCAGCAGCAGCAG 1200
Qy 422 LeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAlaTyr 441
Db 1201 CUCAGCGUCUCCAGAGACCUCCAGAAUUCGGGGAGGAGAAUUCUGACAAUUGGCGCCUAC 1260
Qy 442 SerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGluLeu 461
Db 1261 UCGCUGACCCUGCAAGGCGUGGCAUCAGCUGGCGUGGCGUCUCACUGAGGGAACUG 1320
Qy 462 GlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrValPro 481
Db 1321 GGCAGUGGACUGGCGCCUCAUCCACCAUAACACCCACCUCUCUGUCACACGUGGCC 1380
Qy 482 TrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGlu 501
Db 1381 UGGACACAGCUCUUCGGAACCCGACCAAGCUCUGCUCCACACUGCCAAACCGGCCAGAG 1440
Qy 502 AspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTrp 521
Db 1441 GACGAGUGUGGCGAGGCGCUGGCCUGCCACCGCAGCUGGCGCGCCGAGGSGCUGCGUG 1500
Qy 522 GlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCysVal 541
Db 1501 GGUCCAGGGGCCACCCAGUGUGUACAUCGACGAGCCAGUUCUUCGGGGCCAGGAGUGCGUG 1560
Qy 542 GluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCysLeu 561
Db 1561 GAGGAUUGCCGAGUACUGCAGGGGCUCCCGCAGGAGAUUGUGAAUCCAGGCACUCUGUUUG 1620
Qy 562 ProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGluAla 581
Db 1621 CCGUGCCACCUGAGUGCAGGCCCAAGAUUGGUCUCAGACCUUGUUUGACCGGAGGCU 1680
Qy 582 AspGlnCysValAlaCysAlaHisTyrIleAspProProPheCysValAlaArgCysPro 601
Db 1681 GACCAGUGUGGCGCUGGCCCAUAUAAAGAGACCCUCCUUCUGGUGGCGCGCGCC 1740

Qy 602 SerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGluGly 621
Db 1741 AGCGGUGUGAAACCCUGACCUCUCCAUCAUGCCCAUGGAAAGUUUCCAGAUAGAGGGGC 1800
Qy 622 AlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLysGly 641
Db 1801 GCAUGCCAGCCUUGGCCCAUCAACUGCAGCCACUCCUGUGUGGACCUGGAGACAAGGC 1860
Qy 642 CysProAlaGlu 645
Db 1861 UCCCCCGCCGAG 1872

Search completed: October 16, 2003, 17:20:37
Job time : 351.963 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 872.869 Seconds
(without alignments)
2842.104 Million cell updates/sec

Title: SEQ6

Perfect score: 5078

Sequence: 1 MELAAACRWGLLLALLPPGA.....TFKGTPTAENPEYGLDVPV 919

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame_p2n.model -DEV-xlp
-Q/cyn2_1/USPPO_spool_p/HOLLERAN480/runat_15102003_131912_20526/app_query.fasta_1.4685
-DB=N_Geneseq_19Jun03 -QPMT-fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosom62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=0 -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGNVL_1_1758 -runat_15102003_131912_20526 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4900	96.5	3768	17	AAT40739
2	4900	96.5	3768	20	Human HER-2/neu oncogene
3	4900	96.5	3768	22	Human HER-2/neu pr
4	4900	96.5	3768	24	Human HER-2/neu pr
5	4900	96.5	3768	24	Human HER-2/neu CD
6	4900	96.5	3768	24	Human HER-2/neu DN
7	4900	96.5	4472	21	CDNA encoding the
8	4900	96.5	4473	19	Human tumour anti
9	4900	96.5	4473	20	HER-2 nucleic acid
10	4900	96.5	4473	24	Human gene express
11	4900	96.5	4473	24	Human Her-2 DNA
12	4892	96.3	3768	21	Human heregulin 2
13	4892	96.3	3768	24	Human ERBB2 polyu
14	4892	96.3	3768	24	Human ERBB2 DNA fr
15	4892	96.3	3768	24	Human HER-2 CDNA
16	4892	96.3	3768	24	Human Her2 antigen
17	4892	96.3	3768	24	Human ERBB2 DNA SE
18	4892	96.3	3768	24	Human polynucleoti
19	4892	96.3	3768	24	Human HER2 (ERBB2)
20	4892	96.3	4530	16	Her-2/neu (ERBB2/c
21	4892	96.3	4530	18	Human HER2 gene
22	4892	96.3	4530	21	Nucleotide sequenc
23	4892	96.3	4530	22	Human tyrosine kin
24	4892	96.3	4530	24	Human gene express
25	4892	96.3	4530	24	Breast carcinoma r
26	4892	96.3	4530	24	Human HER2-neu SEQ
27	4892	96.3	4530	24	Human CDNA differe
28	4892	96.3	4530	25	Breast cancer asso
29	4892	96.3	4530	25	ACC03139
30	4892	96.3	9274	24	ABQ83856
31	4892	96.3	9274	24	Human Her2/Neu enc
32	4885	96.2	9274	24	ABK14057
33	4857	95.6	4299	14	HER2 transgene pl
34	4722	93.0	3678	24	HER2 transgene pla
35	4583	90.3	3600	21	Sequence encoding
36	4309.5	84.9	2763	24	Human HER-2/neu co
37	4309.5	84.9	2781	24	Mouse Her-2/neu ex
38	4134	81.4	3955	16	Her-2/neu extracel
39	4134	81.4	3955	21	Rat neu promoter
40	4125.5	81.2	3771	21	Rat Her-2/neu prot
41	4125.5	81.2	3771	22	Mouse Her-2/neu CD
42	4125.5	81.2	3771	24	Nucleotide sequenc
43	3632	71.5	2385	18	Mouse Her-2/neu CD
44	3525	69.4	2871	21	Her2-GM-CSF immuno
45	3422	67.4	1872	11	DC8scFv-erbB2EC fu
					Extracellular port

ALIGNMENTS

RESULT 1
AAT40739
ID AAT40739 standard; cDNA; 3768 BP.
AC
AC AAT40739;
XX
XX
DT 01-JAN-1997 (first entry)
XX
XX HER-2/neu oncogene.

HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
breast cancer; ovary cancer; colon cancer; lung cancer;
prostate cancer; genetic immunisation; tumour; vaccine; vector;
ss.

Homo sapiens.

XX

Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACAGCTCTTTCCGAAACCCACACAGCTCTGCTCCACACTGCAACCGCCA	1500
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGACAGAGTGTGGCGAGGGCTGGCTGCCACCACTGTGCGGCCCGAGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGTCACAGGCCACCCAGTGTGTCACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGGAGTATGTGAATGCCAGGCACTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCGGTGCCACCCCTGAGTGTACGCCCCGAAATGGCTCAGTGACCTGTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys	600
Db	1741	GCTGACCAGTGTGTGGCCTGTGCCACTATAAGGACCTCCCTTCGTGCGTGGCCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTAAACCTGACCTCTCTACATGCCACTCTGGAAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGGCGATGGCAGGCTTGCCCATCAACTGCACCCACTCTCTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGTGCCCCCGCAGAGAGAGCCGCCCTCTGACGTCCATCATCTCGCGTGGTGGC	1980
Qy	653	-----	653
Db	1981	ATTCTGCTGTCGTGCTTTGGGGTGGTCTTTGGGATCCTCATCAAGGCAGCGCAGCAG	2040
Qy	653	-----	653
Db	2041	AAGATCGGAAGTACAGATGCGGAGACTGCTCGAGAAACGGAGCTGTGGAGCCGCTG	2100
Qy	653	-----	653
Db	2101	ACACCTAGCGGAGCGGATGCCCAACCAAGCCACAGATCGGGATCTCTGAAGAGACGGAGCTG	2160
Qy	653	-----	653
Db	2161	AGGAAGGTGAAGTGTCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGCACTCTGGATC	2220
Qy	653	-----	653
Db	2221	CCTGATGGGAGAATGTGAATAATCCAGTGGCCATCAAAAGTGTGAGGGAAAAACATCC	2280
Qy	653	-----	653
Db	2281	CCCAAGCCCAACAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGGGCTCCCCA	2340
Qy	653	-----	653
Db	2341	TATGTCCTCCGCTTCTGGGCATCTGCCTGACATCCAGGCTGCAGCTGGTGACACAGCTT	2400
Qy	653	-----	653
Db	2401	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAAACCGGAGCGCTGGGCTCCCA	2460
Qy	653	-----	653
Db	2461	GACCTGCTGACTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGCGG	2520

QY	653	-----	653
DB	2521	CTCGTACACAGGGACTTTGGCGCTCGGAAGTGTGCTCAAGAGTCCCAACCATGCAA	2580
QY	653	-----	653
DB	2581	ATTACAGACTTCGGGCTGGCTCGCTCGACATTGCACGACAGAGTACCATGCGAGT	2640
QY	653	-----	653
DB	2641	GGGGCAAGTGCCTCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGTTCA	2700
QY	653	-----	653
DB	2701	CACCAGAGTGTATGTGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGG	2760
QY	653	-----	653
DB	2761	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGTGGAAGGGGAGCGG	2820
QY	653	-----	653
DB	2821	CTGCCCGAGCCCCCATCTGCACCATGTGATCTACATGATCATGTGTCAAATCTTGG	2880
QY	653	-----	653
DB	2881	ATTGACTCTGAATGTCGGCCCAAGATTCCGGGAGTTGGTGTCTGAATCTCCGCAT	2940
QY	654	-----GlnAsnGluAspleucglyProAlaSerProLeu	664
DB	2941	AGGACCCCCAGCGCTTTTGGTCAATCCAGAAATGAGGACTTGGGCCCGACGACGT	3000
QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
DB	3001	GACAGCACCTTCTACCGCTCACTGCTGGAGAGCATGACATGGGGGACCTGTGTG	3060
QY	685	GluGlyLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly	704
DB	3061	GAGGAGTATCTGGTTACCCAGCAGCGGCTTCTTGTCTGACAGACCTGCCCGG	3120
QY	705	GlyMetValHisArgHisArgSerSerSerThrArgSerGlyGlyGlyAspLeuThr	724
DB	3121	GGCATGTCTCACACAGCACCGCCAGCTCATCTACAGGAGTGGCGGTGGGACCT	3180
QY	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
DB	3181	CTAGGCGTGGAGCCCTCTGAAGAGGAGGCCCCAGAGTCTCCACTGGCACCTCC	3240
QY	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln	764
DB	3241	GCTGGCTCCGATGTATTTGATGTGACCTGGGAATGGGGCAGCCAAAGGGCTGC	3300
QY	765	LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValPro	784
DB	3301	CTCCCCACATGACCCCGACCCCTCTACAGCGGTACAGTGAGGACCCCAAGTAC	3360
QY	785	ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyr	804
DB	3361	CCCTCTGAGACTGATGGCTACGTTGGCCCCCTGACCTGCAGCCCCCAGCCTCA	3420
QY	805	AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla	824
DB	3421	AACCAAGCAGATCTTCGGCCCCCAGCCCTTCGCCCCGAGAGGGCCCTCTGCCT	3480
QY	825	ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGly	844
DB	3481	CGACCTCTCGTGGTCCCACTCTGGAAGAGGCCCAAGACTCTCTCCCCAGGAAG	3540
QY	845	ValIysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrPro	864
DB	3541	GTCAAAGACGTTTTTGCCTTTGGGGGTGGCGTGGAGAACCCCGAGTACTTGA	3600
QY	865	GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn	884

```

|||||
Db 3601 GGAGGAGTGCCTCAGCCCAACCTCTCTCTCCTTTCAGCCAGCCTTCGACAACCTC 3660
QY 885 TTTTyrTrpAspGlnAspProGluArgGlyAlaProSerThrPheLysGlyThr 904
Db 3661 TATTACTGGACCGACCGACCCAGAGCGGGGGCTCCACCCAGACCTTCAAAGGGACA 3720
QY 905 ProThraGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3721 CCTACGGCAGAGAACCAGAGTACCTGGGTCTGGACGTGCCAGTG 3765

RESULT 2
AA01912
ID AAX01912 standard; DNA; 3768 BP.
XX AC AAX01912;
XX DE
XX DT 21-APR-1999 (first entry)
XX Human HER-2/neu oncogene DNA.
XX HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
XX malignancy; treatment; tumour; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..3768
XX /*tag= a
XX /product= "HER-2/neu"
XX /note= "oncogene"
XX misc_feature 2036..3765
XX /*tag= b
XX /note= "region which elicits immune response"

US5869445-A.
XX
XX 09-FEB-1999.
XX
XX 01-APR-1996; 96US-0625101.
XX
XX 01-APR-1996; 96US-0625101.
XX 17-MAR-1993; 93US-0033644.
XX 12-AUG-1993; 93US-0106112.
XX 31-MAR-1995; 95US-0414417.
XX
XX (UNIW ) UNIV WASHINGTON.
XX
XX Cheever MA, Disis ML;
XX
XX WPI; 1999-152835/13.
XX P-PSDB; AAW92406.
XX
XX Use of HER-2/neu polypeptides - for eliciting an immune response to
XX an HER-2/neu associated malignancy, particularly for treating or
XX preventing tumours
XX
XX Claim 1a; Column 23-32; 26pp; English.
XX
XX This sequence encodes the human HER-2/neu oncogene protein. A fragment
XX of this protein is used in a method for eliciting or enhancing an immune
XX response to HER-2/neu protein. The polypeptide can stimulate T cells and
XX B cells to produce an immune response to the HER-2/neu protein. The
XX method can be used for immunisation against a malignancy in which the
XX HER-2/neu oncogene is associated and in the treatment of an existing
XX tumour, or to prevent tumour occurrence or reoccurrence.
XX
XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 6,12e-227 Length: 3768
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0

```

```

Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 20 Gaps: 1

SEQ6 (1-919) x AAX01912 (1-3768)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGGAGCTGGCGGCTTGTGCCCTGGGGCTCCTCTCGCCCTCTTGCCTCCCGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGACCAAGTGTGCACCGGCACAGACATGAAGCTCGGGCTCCTCTCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACCCAGGGCTGCCAGGTGTGAGGAAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 181 GAACTCACCTACCTGCCCAATGCCAGCTCTCTCTCTCCAGGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTAGCTGCTCATCGCTCAACCAAGTGAGGCAGGTCCCACTGCAGAGGCTCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCCCTGTAGACAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCCTGAACAATACCACTCCCTGTACAGGGGCTCCCCAGGAGGCTCGGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGln 160
Db 421 CAGCTTCGAAGCCTCAGAGATCTTGAAGGAGGGTCTTGATCAGCGGAAACCCCGAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAla 180
Db 481 CTCTGTACCGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACACACAGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPrometCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCGATGTGTAAG 600
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGTGTGGGAGAGAGTTCTGAGGATGTGAGAGCTGACGGCCTGCTGTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCCATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCGCTGCTGCCACTTCAACCAC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGTGAGCTGCACCTGCCAGGCCCTGGTCACTACACACACAGACAGTTTGA 840
QY 281 SerMetProAsnProGluArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCGAGGGCCGGTATACATTCCGGCGCCAGCTGTGTGACTGCTGTCCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACCTACCTTTCTACGGAGCTGGGATCCTGCACCTCTGCTGCCCTGCACAACCAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGAGCCCTGTGCCCGCA 1020

```

```
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1021 GTGTGCTATGCTTGGCGATGAGACACTTGGCAGAGGTGAGGCGAGTTACACGTGCCAAT 1080
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 ATCCAGGAGTTTGCTGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC 1140
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1141 TTTGATGGGACCCAGCCTCCAAACACTGCCCGCTCCAGCCAGACGACTCCAAGTGT 1200
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
Db 1201 GAGACTCGGAAGAGATCACAGGTACCTATACATCTCAGCATGGCCGGACAGCCTGCCT 1260
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTCAGCGTCTTCCAGAACCTGCAAGTAAATCCGGGAGCGAATCTGCACAAATGGCGCC 1320
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerThrPleuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGGAA 1380
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGCAGTGGACTGGCCCTCATCCACATACACCCACCTCTGCTTCGTGCACACGGTG 1440
Qy 481 ProThrAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCTTGGGACCAAGCTCTTTCCGAACCCGACCAAGCTCTGTCCACACTGCGCAACCGGCCA 1500
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGACGAGTGTGGGCGAGGGCTGGCCCTGCCACCAAGCTGTGCGCCCGAGGGCACTGC 1560
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGCTCAGGGCCCCCAGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCGAGTACTGCAGGGCTCCCCAGGAGTATGTGAATGCCAGGCACTGT 1680
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCCGTGCCACCTGAGTGTACGCCCAAGTGGCTCAGTGACCTGTCTTTGGACCCGGAG 1740
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrIleAspProProPheCysValAlaArgCys 600
Db 1741 GCTGACCAGTGTGTGGCCTGTGCCACTATAAGGACCTCCCTTCTGCGTGGCCCGCTGC 1800
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTCGAAGTTTCCAGATGAGGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GGGCGATGCCAGCCTTGCCCATCAACTGCACCCACCTCTGTGTGGACCTGGATGACAAG 1920
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGGTGGTGGC 1980
Qy 653 2000
Db 1981 ATTCGTGTCGTGGTCTTGGGGTGGTCTTTTGGGATCCTCATCAAGCAGCGCAGCAG 2040
Qy 653 2080
Db 2041 AAGATCCGGAAGTACACCATCCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTG 2100
```



```

Db      3181 CTAGGGCTGAGCCCTCTGAAGAGGAGGCCCGAGGCTCCAGCTCCAGCCCTCCGAGGG 3240
QY      745  AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db      3241 GCTGGCTCCGATGATTTGATGGTACCTGGGAATGGGGGAGCCCAAGGGGTGCAAGC 3300
QY      765  LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db      3301 CTCCACACATGACCCAGCCCTCTACAGAGGTACAGTGAAGACCCACAGTACCCTG 3360
QY      785  ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db      3361 CCCTCTGAGACTGATGGCTACGTGGCCCTGACCTGCAGCCGCCAGCTGAATATGTG 3420
QY      805  AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db      3421 AACCAAGCCAGATGTTGGGCCCGAGCCCTTCCGCCGAGAGGGCCCTGCTGCTGCC 3480
QY      825  ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db      3481 CGACCTGCTGGTCCCACTCTGGAAGGCCCAAGACTCTCTCCCGAGGAGATGGGGTC 3540
QY      845  ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db      3541 GTCAAGAGCGTTTTCCTTTGGGGTGGCGTGGAGAACCCGAGTACTTGACACCCAG 3600
QY      865  GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
Db      3601 GGAGGAGCTGCCCTCAGCCCCCAGCCCTCTCTGCTGCTTCCAGCCAGCCCTTCGACACCT 3660
QY      885  TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db      3661 TATTACTGGGACCAAGACCCAGAGCGGGGGCTCCACCCAGCACCTTCAAGGGGACA 3720
QY      905  ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db      3721 CCTACGGCAGAACCAACAGAGTACCTGGGTCTGGAGTGGCCAGTG 3765

RESULT 3
AAH23392
ID      AAH23392 standard; DNA; 3768 BP.
AC      AAH23392;
DT      25-SEP-2001 (first entry)
DE      Human HER-2/neu protein encoding DNA.
KW      Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
KW      oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.
OS      Homo sapiens.
XX      Key
FH      Location/Qualifiers
FT      1..3768
FT      /*tag= a
FT      /product= "HER-2/neu protein"
XX      WO200153463-A2.
XX      26-JUL-2001.
XX      19-JAN-2001; 2001WO-US01850.
XX      21-JAN-2000; 2000US-0177545.
PR      (CORI-) CORIXA CORP.
XX      Cheever MA, Hand-Zimmermann S;
XX      WPI: 2001-476112/51.
DR      P-PSDB; AAB85458.

```

```

XX      New antigen-presenting cells, useful as vaccines for eliciting or
PT      enhancing an immune response to HER-2/neu protein, particularly useful
PT      for treating or preventing cancer, e.g. breast cancer
XX      Claim 1; Page 41-46; 49pp; English.
XX      The invention provides an isolated antigen-presenting cell, which
CC      expresses at least an immunogenic portion of a polypeptide that produces
CC      an immune response to HER-2/neu protein. The antigen-presenting cells are
CC      useful as vaccines for eliciting or enhancing an immune response to
CC      HER-2/neu protein, particularly in treating or preventing malignancies in
CC      which the HER-2/neu oncogene is associated. Specifically, these are
CC      useful for treating or preventing cancer, e.g. breast cancer, ovarian,
CC      colon, lung or prostate cancers. The present sequence represents a DNA
CC      encoding the human HER-2/neu protein (also known as p185 or c-erbB2).
XX      Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
SQ

```

```

Alignment Scores:
Pred. No.: 6,12e-227 Length: 3768
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 22 Gaps: 1

```

```

SEQ6 (1-919) x AAH23392 (1-3768)
QY      1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProGlyAla 20
Db      1 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCTCGCCCTCTGCCCCGAGGC 60
QY      21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProLaserProGlu 40
Db      61 GCGAGCACCAAGTGTGCACCGGCACAGACATCAAGCTCGCGCTCCCTGCCAGTCCCGAG 120
QY      41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db      121 ACCCAGCTGGACATGCTCGCCACCTCTACAGGCTGCGAGGTGGTGGAGGAAACCTG 180
QY      61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspGlyGlnVal 80
Db      181 GAACCTACCTAGCTGCTCCCAATGCCAGCTGTCTCTCTCGAGGATATCCAGGAGGTG 240
QY      81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      241 CAGGGCTAGTGTCTATCGCTCACCAAGTGAGGCAGGTCCCACTGCAGAGGCTGCGG 300
QY      101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      301 ATGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTAGACATGGA 360
QY      121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      361 GACCCGCTGAACAATACCACTGTGTACAGGGGCTCCCGAGGAGGCTGCGGGAGCTG 420
QY      141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db      421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTATTCAGCGGAACCCCGAG 480
QY      161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAla 180
Db      481 CTCTGCTACCGACACAGATTTTGTGAAGGACATCTTCCACAAGAACCAACAGCTGGCT 540
QY      181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      541 CTCACACTGATAGACCAACCGCTCTCGGGCTGCCACCCCTGTGTCTCCGATGTGAAG 600
QY      201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      601 GGCTCCCGCTGCTGGGGAGAGATTGTGAGGATTGTGAGAGCTGACGCGCTGCTGTGT 660

```


Db	61	CGGAGCACC	AAAGTGTG	CACCGGCAC	ACATGAAG	CTGGCGCT	CCCTGCGC	AGTCCCGAG	120
Qy	41	ThrHisLeu	AspMetLeu	ArgHisLeu	TyrGlnGly	CysGlnVal	ValGlnGly	AsnLeu	60
Db	121	ACCCACCT	TGGACATG	CTCCGCCAC	CTCTACCA	GGGCTGCC	AGGTGGT	GCAGGAAAC	180
Qy	61	GluLeuThr	TyrLeuPro	ThrAsnAla	SerLeuSer	PheLeuGln	AspIleGln	GluVal	80
Db	181	GAAC	TCACCT	ACCTGCGC	CAACCAAT	GCCAGCGT	CTCTTC	CTGCAAGATAT	240
Qy	81	GlnGlyTyr	ValLeuIle	AlaHisAsn	GlnValArg	GlnValPro	LeuGlnArg	LeuArg	100
Db	241	CAGG	CTACGCT	TCATCGT	CTACAAC	CAAGTAGG	CAGGCTCC	CACTGCAGAGCT	300
Qy	101	IleValArg	GlyThrGln	LeuPheGlu	AspAsnTyr	AlaLeuAla	ValLeuAsp	AsnGly	120
Db	301	ATTGTG	CGAGC	ACCCAGCT	TTTGAGG	ACAACTAT	GCCCTGG	CCGTGTAGACAA	360
Qy	121	AspProLeu	AsnAsnThr	ThrProVal	ThrGlyAla	SerProGly	GlyLeuArg	GluLeu	140
Db	361	GACC	CGCTGA	ACAATACC	ACCCCTG	TACAGG	GGCCTCCC	CAGGAGCCT	420
Qy	141	GlnLeuArg	SerLeuThr	GluIleLeu	LysGlyVal	LeuIleGln	ArgAsnPro	Gln	160
Db	421	CAGCTT	CGAACCT	TCACAGAT	CTTGAAG	AGGGGTCT	TGATCC	AGCGGAAC	480
Qy	161	LeuCysTyr	GlnAspThr	IleLeuTrp	LysAspIle	PheHisLys	AsnAsnGln	LeuAla	180
Db	481	CTCTGCT	ACCAGG	ACAGATTT	TGTGAAG	GACATCTT	CCAAAGA	CAACCACTGGCT	540
Qy	181	LeuThrLeu	IleAspThr	AsnArgSer	ArgAlaCys	HisProCys	SerProMet	CysLys	200
Db	541	CTCACAC	TGATAG	ACACCA	CCGCTCT	CGGSCCT	GCCACCC	CTGTTCTCG	600
Qy	201	GlySerArg	CysTrpGly	LeuSerSer	GluAspCys	GlnSerLeu	ThrArgThr	ValCys	220
Db	601	GGCT	CCCGCT	GCTGGG	GAGAGTCT	TGAGGAT	TGTCAG	AGCCTGTCTGT	660
Qy	221	AlaGlyLys	CysAlaArg	CysLysGly	ProLeuPro	ThrAspCys	CysHisGlu	GlnCys	240
Db	661	GCCGGT	GGCTGT	GCCCTG	CAAGGG	CCACTG	CCCACT	GACTGCTGCC	720
Qy	241	AlaAlaGly	CysThrGly	ProLysHis	SerAspCys	LeuAlaCys	LeuHisPhe	AsnHis	260
Db	721	GCTGCC	GGCTG	CACGGC	CCCCAAC	CACTCTG	ACTGCTG	CGCTCC	780
Qy	261	SerGlyLys	CysGluLeu	HisCysPro	AlaLeuVal	ThrTyrAsn	ThrAspThr	PheGlu	280
Db	781	AGTGGC	ATCTGT	GAGCTG	CACTGCC	AGCCCTG	CTACCT	TACAA	840
Qy	281	SerMetPro	AsnProGlu	GlyArgTyr	ThrPheGly	AlaSerCys	ValThrAla	CysPro	300
Db	841	TCCATG	CCCCAAT	CCGAGG	CGGTAT	TACATTC	GGCGCC	AGCTGTG	900
Qy	301	TyrAsnTyr	LeuSerThr	AspValGly	SerCysThr	LeuValCys	ProLeuHis	AsnGln	320
Db	901	TACAAC	TACCTTT	CTACGG	AGTGGG	ATCTG	CACTG	CTGCTGCC	960
Qy	321	GluValThr	AlaGluAsp	GlyThrGln	ArgCysGlu	LysCysSer	LysProCys	AlaArg	340
Db	961	GAGGTG	CAGCAG	AGGATG	GAACAC	AGCGGT	TGTGAG	AAGTGC	1020
Qy	341	ValCysTyr	GlyLeuGly	MetGluHis	LeuArgGlu	ValArgAla	ValThrSer	AlaAsn	360
Db	1021	GTGTGCT	ATGTCT	TGGG	CACTTGG	AGAGGT	GAGG	CAGTTAC	1080
Qy	361	IleGlnGlu	PheAlaGly	CysLysLys	IlePheGly	SerLeuAla	PheLeuPro	GluSer	380
Db	1081	ATCCAG	GAGTTG	CTGCTG	CAACA	GATCTTT	GGGAGC	CTGGCATTT	1140
Qy	381	PheAspGly	AspProAla	SerAsnThr	AlaProLeu	GlnProGlu	GlnLeuGln	ValPhe	400
Db	1141	TTTGAT	GGGACC	CAGCTC	CAACACT	CGCCCG	CTCCAG	CCAGAGC	1200
Qy	401	GluThrLeu	GluGluIle	ThrGlyTyr	LeuTyrIle	SerAlaTrp	ProAspSer	LeuPro	420
Db	1201	GAGACT	CTGG	NAGAGAT	CACAGTT	TACTAT	ATCTC	AGCATG	1260
Qy	421	AspLeuSer	ValPheGln	AsnLeuGln	ValIleArg	GlyArgIle	LeuHisAsn	GlyAla	440
Db	1261	GACCTC	AGCGTCT	TCCAGA	ACCTGCA	AGTAAT	CCGGG	ACCAATTC	1320
Qy	441	TyrSerLeu	ThrLeuGln	GlyLeuGly	IleSerTrp	LeuGlyLeu	ArgSerLeu	ArgGlu	460
Db	1321	TACT	CGCTG	ACCTG	CAAGGCT	GGCATC	AGCTGG	CTGGGCT	1380
Qy	461	LeuGlySer	GlyLeuAla	LeuIleHis	AsnThrHis	LeuCysPhe	ValHisThr	Val	480
Db	1381	CTGG	GCAGT	GGACTG	GGCCTC	ATCACC	ATPAAC	CAACCACT	1440
Qy	481	ProTrpAsp	GlnLeuPhe	ArgAsnPro	HisGlnAla	LeuLeuHis	ThrAlaAsn	ArgPro	500
Db	1441	CCCT	GGGACC	AGCTCT	TTTTCG	GAACCC	CAAGCTCT	GTCTCC	1500
Qy	501	GluAspGlu	CysValGly	GlyGluGly	LeuAlaCys	HisGlnLeu	CysAlaArg	GlyHisCys	520
Db	1501	GAGG	ACGAGT	GTGTGG	CGGAGG	CGCTGG	CCCTGCC	CAAGCTGT	1560
Qy	521	TrpGlyPro	GlyProThr	GlnCysVal	AsnCysSer	GlnPheLeu	ArgGlyGln	GluCys	540
Db	1561	TGGG	GTCCAG	GGCCAC	CCAGTGTCT	CACTGC	AGCAGCTTC	TTCTCG	1620
Qy	541	ValGluGlu	CysArgVal	LeuGlnGly	LeuProArg	GlyTyrVal	AsnAlaArg	HisCys	560
Db	1621	GTGG	AGGAAT	CGCCAG	TACTGC	AGGGCT	CTCCC	AGGAGTAT	1680
Qy	561	LeuProCys	HisProGlu	CysGlnPro	GlnAsnGly	SerValThr	CysPheGly	ProGlu	580
Db	1681	TTGCG	GTGCC	ACCCTG	AGTGTG	CAAGCC	AGAAATGG	CTGCTG	1740
Qy	581	AlaAspGln	CysValAla	CysAlaHis	TyrLysAsp	ProPheCys	ValAlaArg	Cys	600
Db	1741	GCTG	ACCACTGT	TGGGCTGT	GCCCACTAT	AAGACCC	CTCCCTCT	CTCGG	1800
Qy	601	ProSerGly	ValLysPro	AspLeuSer	TyrMetPro	IleTrpLys	PheProAsp	GluGlu	620
Db	1801	CCCAG	CGGTGT	GAACCTG	ACCTCTC	TCTACAT	GCCATCTG	GAAATTT	1860
Qy	621	GlyAlaCys	GlnProCys	ProIleAsn	CysThrHis	SerCysVal	AspLeuAsp	AspLys	640
Db	1861	GGCG	CATGCC	AGCCTT	GCCCCAT	CACTGCA	CCCCACT	CTCTGT	1920
Qy	641	GlyCysPro	AlaGluGln	ArgAlaSer	ProLeuThr	Ser			653
Db	1921	GGTG	CCCCCG	CGAGCAG	AGCCGCTCT	GACGTCC	ATCATCT	CTGCGG	1980
Qy	653								653
Db	1981	ATTCT	GTGCTG	TGCTG	TTGGG	GTGTTCT	TTGGG	ATCCTCAT	2040
Qy	653								653
Db	2041	AAGAT	CCGGAAG	TACAGAT	GCAGAG	ACTGCTG	CAGGAA	ACGGAGCT	2100
Qy	653								653
Db	2101	ACAC	CTAGC	GAGG	AGTGC	CAAC	AGCGC	AGATCG	2160
Qy	653								653
Db	2161	AGGA	AGGTGA	AGGTG	CTTGG	ATCTT	TGGC	GTCTACA	2220
Qy	653								653
Db	2221	CTGT	ATGGG	GAGA	GTGAAA	ATTTCC	AGTGCC	ATCAAG	2280

QY 653 ----- 653
 Db 2281 CCCAAAGCCAAACAAAGAAATCTTAGAGAAAGCATACGTGATGCTGGTGGCTCCCA 2340
 QY 653 ----- 653
 Db 2341 TATGTCTCCGCTTCTGGGCATCTGCTCATCCACGGTGACGCTGACACAGCTT 2400
 QY 653 ----- 653
 Db 2401 ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGCGGAGCCCTGGGCTCCAG 2460
 QY 653 ----- 653
 Db 2461 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGG 2520
 QY 653 ----- 653
 Db 2521 CTGCTACACAGGAGCTTGGCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA 2580
 QY 653 ----- 653
 Db 2581 ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGACAGAGTACCATGCAGAT 2640
 QY 653 ----- 653
 Db 2641 GGGGCAAGTGCCCATCAAGTGGATGGCTGGAGTCCATTCTCCGCGCGGCTTCACC 2700
 QY 653 ----- 653
 Db 2701 CACCAGAGTATGTGTGGAGTTATGTGTGACTGTGTGGAGGCTGATGACTTTTGGGGCC 2760
 QY 653 ----- 653
 Db 2761 AAACCTTACATGGGATCCAGCCCGGAGATCCCTGACCTGTCTGAAAGGGGAGCGG 2820
 QY 653 ----- 653
 Db 2821 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTCAAATGTTGGATG 2880
 QY 653 ----- 653
 Db 2881 ATTGACTCTGAATGTCGCCAAGATTCCGGGAGTTGGTGCTGTGAATTCCTCCGATGGCC 2940
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 Db 2941 AGGGACCCCGAGCGCTTTGTGGTCAATCCAGATGAGGACTTGGGCCACCGCCTGCTGG 3000
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 Db 3001 GACAGCACCTTCTPACCGCTCACCTGTGGAGGACGATGACATGGGGACCTGGTGATGCT 3060
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
 Db 3061 GAGGAGTATCTGTATCCAGCCAGCAGGCGCTTCTGTCCAGACCTGCCCGCGGCGCTGG 3120
 QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
 Db 3121 GGCTGGTCCACCACAGGACCCGACGCTCATCTACCAGGAGTGGCGTGGGACCTTGACA 3180
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 Db 3181 CTAGGGCTGGAGCCCTCTAGAGAGAGGCCCCCAGGCTCTCCACTGGCACCTCCGAGGG 3240
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
 Db 3241 GCTGGCTCCGATGATTTGATGTCAGCTGGGAATGGGGACGCCAAGGGGCTGCAAGC 3300
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerClnuAspProThrValProLeu 784
 Db 3301 CTCCCCACACATGACCCCGCCCTCTACAGCGGTACAGTGAGGACCCACAGTACCCCTG 3360
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804

Db 3361 CCTCTGAGACTGATGGCTAGCTTCCCTCCCTGACCTGAGCCCTGACCTGAATATGTG 3420
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
 Db 3421 AACCAAGCCAGATGTTTGGCCGCCAGCCCTTCCGCCCGAGAGGGCCCTCTGCTGTGCC 3480
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
 Db 3481 CGACCTGCTGGTGGCCACTCTGGAAAGGCCCAAGACTCTCTCCAGGGAGAAATGGGGTC 3540
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
 Db 3541 GTCAAAGACGTTTTTCCCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3600
 QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3601 GGAGGAGCTGCCCTCAGCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3660
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
 Db 3661 TATTACTGGGACGAGACCCACAGAGCGGGGGGCTCCACCCAGACCTTCAAGAGGACA 3720
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3721 CCTACGGCAGAGAAACCCAGAGTACCTGGGTCTGGACCTGCCAGTG 3765
 RESULT 5
 ID ABA92250 standard; cDNA; 3768 BP.
 AC ABA92250;
 DT 17-JUN-2002 (first entry)
 DE Human Her-2/neu cDNA.
 KW Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
 KW receptor; human; gene therapy; gene; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX CDS 1..3768
 XX FT /*tag= a
 XX FT /product= "Her-2/neu"
 XX PN W0200212341-A2.
 XX PD 14-FEB-2002.
 XX PF 03-AUG-2001; 2001WO-US24283.
 XX PR 03-AUG-2000; 2000US-0632507.
 XX PA (CORI-) CORIXA CORP.
 XX PA (SMIX) SMITHKLINE BEECHAM BIOLOGICALS.
 XX PI Cheever MA, Gheysen D;
 XX WIPI; 2002-241743/29.
 XX P-PSDB; AAW51143.
 XX Her-2/neu fusion protein for treating or preventing cancer by eliciting
 PT or enhancing an immune response to the protein, has Her-2/neu
 PT extracellular domain fused to Her-2/neu intracellular or
 PT phosphorylation domain
 XX Disclosure; Fig 15; 141pp; English.
 PS The present sequence is that of human Her-2/neu oncogene cDNA.
 CC The cDNA encodes Her-2/neu (p185), an oncogenic self protein and
 CC target for anti-cancer vaccines. The Her-2/neu gene is amplified

CC and p105 is overexpressed in a variety of cancers, including breast,
 CC ovarian, colon, lung and prostate cancer. Her-2/neu is a member
 CC of the tyrosine kinase family of receptor-like glycoproteins.
 CC Its overexpression correlates with a poor prognosis in breast and
 CC ovarian cancers. The invention provides Her-2/neu fusion
 CC proteins, nucleic acids encoding them, viral vectors, and vaccines
 CC comprising the fusion proteins or nucleic acid molecules. In
 CC preferred fusion proteins, the extracellular domain of a Her-2/neu
 CC protein is fused to a Her-2/neu intracellular domain or
 CC phosphorylation domain (or its Deltapp fragment). An immune
 CC response to Her-2/neu protein is elicited or enhanced by
 CC administering the fusion protein in the form of a vaccine, or by
 CC transfecting cells of an animal *ex vivo* with a nucleic acid
 CC encoding the fusion protein, and delivering the transfected cells
 CC to the animal. The fusion proteins, nucleic acids, and isolated
 CC specific T-cells are useful for inhibiting the development of a
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer
 CC in a patient. T cells that specifically react with a Her-2/neu
 CC fusion protein can be used to remove tumour cells from a sample in
 CC order to inhibit the development of cancer in a patient.
 XX
 SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:

Pred. No.: 6.12e-227 Length: 3768
 Score: 4900.00 Matches: 919
 Percent Similarity: 73.23% Conservative: 0
 Best Local Similarity: 73.23% Mismatches: 0
 Query Match: 96.49% Indels: 336
 DB: 24 Gaps: 1

SE06 (1-919) x ABA92250 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 DB 1 ATGGAGCTGGCGGCTTGGCGCTGGGGGCTCTCTCCGCGCTTGGCCCCGGAGGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 61 GCGAGCACCAAGTGTGCACCGGCACACATGAAGCTGCGCTCCCTGCCAGCTCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 DB 121 ACCCACTGGACATGCTCCGCGACGCTCTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTG 180
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 DB 181 GAACCTACCTACCTGCCCCCACCAGTCCAGCTGCTCTCTCCAGGATATCCAGAGGTG 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 241 CAGGCTACGTGCTCATCGCTCACAACTGAGGAGGAGTCCCACTGCAGAGCTGGG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCGTGTAGACAAATGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB 361 GACCCGCTGAACAATACCCACCCCTGTACAGGGGCTCCCGAGAGGCTTCGGGAGCTG 420
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
 DB 421 CAGCTTGAAGCCCTCACAGATCTTGAAGAGGAGGCTTTGATCCAGGGAACCCCGAG 480
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 DB 481 CTCTGCTACCAAGGACACGATTTTGTGAAGAGACATCTTCCACAAGAACCAACAGCTG 540
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 541 CTCACATGATAGACACCAACCGCTCTGGGGCTTGCACCCCTGTTCTCCGATGTGAAG 600

QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 601 GGCTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGCGCAGCTGCTGT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 DB 661 GCCGGTGGCTGTGCCCGCTGCAAGGGGCACTGCCCACTGACTGTGCTGCATGAGCAGTGT 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 721 GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCTGGCGCTGCTCCACTTCAACCAAC 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 781 AGTGCATCTGTGAGCTGCACTGCCAGCCCTGGTGCACCTACAACACAGACACAGTTTGA 840
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 841 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCAGCTGTGTGACTGCTGCTGCC 900
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 901 TACAACCTACCTTTCTACGGAGCTGGGATCTTCACCCCTCGTCTGCCCTTGCACCAACCA 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAGAGTGCAGCAAGCCCTGTGCCCGA 1020
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1021 GTGTGCTATGCTTGGCATGGAGCATTGGCAGAGGTGAGGGCAGTTTACCAGTGCCTAAT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1081 ATCCAGGAGTTTGGCTGCAAGAAGATCTTTGGAGCGCTGGCAATTTTTCGGGAGAGC 1140
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB 1141 TTTGATGGGACCCAGCCCTCCACACACTGCCCGCTCCAGCCAGCAGCAGCTCCAAGTGT 1200
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro 420
 DB 1201 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGTCGCGCAGCAGCTGCCT 1260
 QY 421 AspleuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1261 GACCTCAGCTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTTCTGCACAAATGGCGCC 1320
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1321 TACTCGCTGACCTTGAAGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGGA 1380
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1381 CTGGGCACTGAGCTGGCCCTCATCACCATTAACCCCACTCTGCTTGTGCTGCACACGCTG 1440
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1441 CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCCAACCCGGCCA 1500
 QY 501 GluAspGluCysValGlyGluAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1501 GAGGACGAGTGTGGGCGAGGGCTGGCCCTGCCACAGCTGTGCGCCCGAGGCGCACTGC 1560
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1561 TGGGGTCCAGGGCCCACTGCTCAACTGTCAGCCAGCTTCTTCTCGGGGCGCAGGAGTGC 1620
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaAatqHisCys 560
 DB 1621 GTGGAGGAATGCCAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT 1680
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580

[illegible]

	Db	 TCCATGCCCAATCCCGAGGCGGGTACATACTCGGGCCAGCTGTGTGACTGCCTGTCCC	1074
	Qy	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
	Db	 TACAACCTACCTTTCTACGGACGTGGGATCTCGACCCTCGTCGTGCCCTTGCAACACAA	1134
	Qy	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
	Db	 GAGGTGACACGAGGATGGAACACACACGCGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1194
	Qy	ValCystTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
	Db	 GTGTGCTATTGCTGGCATGTGGACACTTCGCGAGAGGTGAGGCGAGTTACCACTGCCAAT	1254
	Qy	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
	Db	 ATCCAGAGATTGCTGGCTCGAAGAAGATCTTTGGAGCTTGCATTTCTGCCGGAGAGC	1314
	Qy	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
	Db	 TTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1374
	Qy	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
	Db	 GAGACTCTGAAGAGATCACAGGTACCTATACATACTACATGTGGCGGCGGACACGCTGCCT	1434
	Qy	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
	Db	 GACCTACGCTCTCCAGAACCTGCAGTAGTAATCCGGGAGCAATCTGCAACAATGGCGCC	1494
	Qy	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
	Db	 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGGCTGCCTCACTGAGGAA	1554
	Qy	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
	Db	 CTGGCGAGTGGACTGGCCCCATCTCCACATPAACACCCACCTCTGCTGTGTCACACGGTG	1614
	Qy	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
	Db	 CCCTGGGACAGCTCTTTCGGAACCCGCCACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1674
	Qy	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
	Db	 GAGCACAGTGTGTGGCGAGGGCTTGGCTGCCACCCAGCTGTGCGCCCGAGGCGACTGC	1734
	Qy	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
	Db	 TGGSGTCCAGGGCCACCCAGTGTCTCACTGCAGCCAGTTCCTTCTGGGGCCAGGAGTGC	1794
	Qy	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
	Db	 GTGAGGAATGCCAGTACTTGCAGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1854
	Qy	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
	Db	 TTGCCGTGCCACCTTGAGTCTCAGTCCAGAACATGGCTCAGTGACCTGTTTGGACCGGAG	1914
	Qy	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys	600
	Db	 GCTGACCAAGTGTGGCTGTGCCACTTGCCCACTATAAAGGACCCCTCCCTTCTCGTGGCCCGCTGC	1974
	Qy	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
	Db	 CCAGCGGTGTGAACCTGACCTTCTCATGCCCATCTGGAAGTTCCAGATGAGGAG	2034
	Qy	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
	Db	 GGCGCATGCCAGCTTGGCCCATCACTGCACCCACTCTCTGTGTGNACCTGGATGACAAG	2094
	Qy	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653

2095	Db	GGCTGCCCCCGCAGCAGAGAGCCGCTCTGACGTCCATCATCTCTGCGGTGGTTGGC	2155
653	QY	-----	653
2155	Db	ATTCTGCTGTCGTGGGTCTTTGGGGTGGTCTTTTGGGATCCTTCATCAAGCGACGCAGCAG	2214
653	QY	-----	653
2215	Db	AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAACGGAGCTGTTGGAGCCGCTG	2274
653	QY	-----	653
2275	Db	ACACCTAGCGGAGCGATGCCACACAGCGCAGATCGGGAATCCTGAAAGAGACGGAGCTG	2334
653	QY	-----	653
2335	Db	AGGAAGGTGAAGGTGCTTGGATCTGGGGCTTTTGGGCACAGCTCTACAAGGCATCTGGATC	2394
653	QY	-----	653
2395	Db	CCTGATGGGGAGAAATGTGAAAAATTCAGTGGCCATCAAAGTGTGGAGGAAAAACACATCC	2454
653	QY	-----	653
2455	Db	CCAAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGTGGCTCCCCA	2514
653	QY	-----	653
2515	Db	TATGTCCTCCGCCCTTCTGGGCATCTGCCTCACATCCACGGTGCAGCTGTTGTGACACAGCTT	2574
653	QY	-----	653
2575	Db	ATGCCCTATGGCTGCCTCTTAGACCATGTCCTGGGAAAAACGCGGACGCTGGCTCCCG	2634
653	QY	-----	653
2635	Db	GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGCGG	2694
653	QY	-----	653
2695	Db	CTCGTACAGAGGACTTGGCCGCTCGGAAGTGCTGGTCAAGAGTCCCAACCATGTCAAA	2754
653	QY	-----	653
2755	Db	ATTACAGACTTCGGCTGGCTGCTGGACATTGACGAGACAGAGTACCATGCAGAT	2814
653	QY	-----	653
2815	Db	GGGGCAAGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCCGCGCGTTCAAC	2874
653	QY	-----	653
2875	Db	CACCAGAGTGATGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC	2934
653	QY	-----	653
2935	Db	AAACCTTACGATGGATCCCGACCCCGGAGATCCCTGCACCTGCTGGAAGGGGAGCGG	2994
653	QY	-----	653
2995	Db	CTGCCCGACGCCCATCTGCACCATTTGATGTCTACATGATCATGTGTCAAATGTTGGATG	3054
653	QY	-----	653
3055	Db	ATTGACTCTGAATGTCGGCCCAAGATTCCGGGAGTTGGTGTCTGAATTCCTCCCATGGCC	3114
654	QY	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
3115	Db	AGGGACCCCGCGCTTTGTGGTTCATCCAGATGAGGACTTGGGCCCGACCCAGTCCCTTG	3174
665	QY	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
3175	Db	GACAGACCTTCTACCGCTCACCTGCTGGAGAGCAGTACATGGGGACCTGGTGGATGCT	3234

Qy	685	Gluc	Ty	r	Leu	Val	Pro	Gln	Gln	Gly	Phe	Phe	Cys	Pro	Asp	Pro	Ala	Pro	Gly	Ala	Gly	704	
Db	3235	GAG	AG	TAT	CT	GT	TAC	CCC	AG	CAG	GGC	TCT	TCT	TG	CC	AG	ACC	CTG	CCCC	GG	CG	3294	
Qy	705	Gly	Met	Val	His	His	Arg	His	Arg	Ser	Ser	Ser	Thr	Arg	Ser	Gly	Gly	Gly	Arg	Asp	Leu	Thr	724
Db	3295	GGC	AT	GGT	CC	CAC	CAG	CAC	CCG	CAG	CT	CAT	C	TAC	CAG	AG	TAG	TGG	CGT	GGG	AC	3354	
Qy	725	Leu	Gly	Leu	Glu	Pro	Ser	Glu	Glu	Glu	Ala	Pro	Arg	Ser	Pro	Leu	Ala	Pro	Ser	Glu	Gly	744	
Db	3355	CTA	GGG	CT	GAG	CCC	TCT	GA	AG	AGG	AGG	CCC	CAG	GGT	CTC	ACT	TGG	CAC	CTC	CG	AA	3414	
Qy	745	Ala	Gly	Ser	Asp	Val	Phe	Asp	Gly	Asp	Leu	Gly	Met	Cly	Ala	Ala	Lys	Gly	Leu	Gln	Ser	764	
Db	3415	GC	T	GGT	TCC	AT	G	TAT	T	GT	AT	G	T	AG	T	CT	GGG	AG	CC	AA	GG	3474	
Qy	765	Leu	Pro	Thr	His	Asp	Pro	Ser	Pro	Leu	Gln	Arg	Tyr	Ser	Glu	Asp	Pro	Thr	Val	Pro	Leu	784	
Db	3475	CT	CCC	CAC	ACAT	GAC	CCC	ACG	CCC	TCT	PAC	ACG	GGT	TAC	AGT	CAG	GAC	CCC	ACC	CAG	TAC	3534	
Qy	785	Pro	Ser	Glu	Thr	Asp	Gly	Tyr	Val	Ala	Pro	Leu	Thr	Cys	Ser	Pro	Gln	Pro	Glu	Tyr	Val	804	
Db	3535	CC	CT	CT	GAG	ACT	GAT	GG	CT	ACG	T	TG	CCC	CT	GAC	CT	GAC	CCC	AG	CCC	AG	3594	
Qy	805	Asn	Gln	Pro	Asp	Val	Arg	Pro	Gln	Pro	Pro	Ser	Pro	Ser	Pro	Arg	Gly	Pro	Leu	Pro	Ala	824	
Db	3595	AAC	CAG	CC	AGAT	GT	TCG	CCCC	ACG	CCCC	CTT	CG	CCC	CGAG	AGG	CC	CTC	TGC	CT	GCT	GC	3654	
Qy	825	Arg	Pro	Ala	Gly	Ala	Thr	Leu	Glu	Arg	Pro	Lys	Thr	Leu	Ser	Pro	Gly	Lys	Asn	Gly	Val	844	
Db	3655	CG	ACC	TG	CT	GGT	CC	ACT	CT	GGA	AA	GCCC	CA	AG	ACT	CT	C	CCC	CAG	GA	AA	3714	
Qy	845	Val	Lys	Asp	Val	Phe	Ala	Phe	Gly	Gly	Ala	Val	I	Asn	Pro	Glu	Tyr	Leu	Thr	Pro	Gln	864	
Db	3715	GT	CAA	AG	ACG	CTT	T	TG	CTT	TGG	GGT	TG	CCG	TG	GAG	AA	CCCC	GAG	TACT	TG	AC	3774	
Qy	865	Gly	Gly	Ala	Ala	Pro	Gln	Pro	His	Pro	Pro	Pro	Ala	Phe	Ser	Pro	Ala	Phe	Asp	Asn	Leu	884	
Db	3775	GG	AG	GAG	CT	GC	CC	CT	CAG	CCCC	ACC	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	3834	
Qy	885	Tyr	Tyr	Trp	Asp	Gln	Asp	Pro	Pro	Glu	Arg	Gly	Ala	Pro	Ser	Thr	Phe	Lys	Gly	Thr		904	
Db	3835	TAT	T	ACT	TGG	AC	CAG	GC	CC	ACC	CAC	CAG	CGG	GGG	GGT	CTC	CC	CC	CAG	CA	CTT	3894	
Qy	905	Pro	Thr	Ala	Glu	Asn	Pro	Glu	Tyr	Leu	Gly	Leu	Asp	Val	Pro	Val						919	
Db	3895	CT	TAC	GC	CAG	CA	AA	CC	CAG	AGT	ACT	TGG	GT	CT	GG	AGT	GC	CA	GTG	CC	AGT	3939	

DESCRITTORE

RESOL 0
ABQ76220
ID ABQ76220 standard; DNA; 4473 BP.

AC ABO76220:

21-OCT-2002 (first entry)

XX. DE. Human tumour antigen ErbB-2 DNA.

XX Tumour antigen; human; vaccine; cellular immune response; immunogen;
KW cancer; tumour; ErbB-2; ds.

OS Homo sapiens.

US6287569-B1.

XX
PD
11-SEP-2001XX
FD
06-APR-1968:XX
DB 10-3 DB-1007-0700-0424678[illegible]

XX	Kipps TJ, Wu Y;
XX	
XX	WPI; 1998-583198/49.
XX	
XX	Generating cellular immune response in patient to target protein -
PT	comprises introducing vector with nucleotide sequence encoding
PT	immunogen comprising protein processing signal into cell of patient
XX	
XX	Disclosure; Column 81-86; 61pp; English.
XX	
CC	This invention describes a novel method for generating a cellular immune
CC	response in a patient to a target protein or its fragment. The method
CC	involves introducing a vector containing a nucleotide sequence encoding
CC	a chimeric immunogen comprising a protein processing signal and the
CC	target protein or its fragment. The immunogen is produced by the cells
CC	and processed so that the target protein or its fragment is presented to
CC	the patients immune system and a cellular immune response is initiated.
CC	The method and vectors can be used as a form of vaccination and could be
CC	used to generate a cellular immune response in patients to, e.g.
CC	cancerous tumours. The cellular immune response is the predominant immune
CC	response in the patient. This sequence represents a DNA fragment which
CC	encodes the human tumour antigen ErbB-2 described in the method of the
CC	invention.
CC	
CC	Note: The information in this spec has been previously disclosed in
CC	WO19845444 however this spec contained no sequence information.
XX	
XX	Sequence 4773 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other;

Alignment Scores:	7,31e-227	Length:	4473
Pred. No.:	Score:	Matches:	919
	4900.00	Conservative:	0
Percent Similarity:	73.23%	Mismatches:	0
Best Local Similarity:	73.23%	Indels:	336
Query Match:	96.49%	Gaps:	1
DB:			

SEQ6 (1-919) x ABQ76220 (1-4473)

QY	1	MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla	20
DB	175	ATGGAGCTGGCGGCTTGTGGCGCTGGGGGCTCTCTCTCGCCCTTTGCCCGCCGAGCC	234
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
DB	235	GGGAGCACCAGTGTGACCGGCACACATGAAGCTGGCGCTCCCTGCCAGTCCCGAG	294
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu	60
DB	295	ACCCACCTGGACATCTCCGCCACCTCTACCAGGCGTCCACAGGTGGTGCGAGGAACCTG	354
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
DB	355	GAACATCACTACTCTGCCACCAATGCCAGCCTGTCTTCTTGCAGGATATCCAGGAGGTG	414
QY	81	GlnGlyTyrValIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
DB	415	CAGGCGCTACGTGCTCATCGCTCACAACTGAGGAGGTCCTCCACATGCAGAGGCTCGGG	474
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
DB	475	ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACATATGCCCTGGCCGTGCTAGACAATGGA	534
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
DB	535	GACCCGCTGAACAATACCACCCCTGTACAGAGGGCGTCCCCAGGAGGCGCTGCGGGAGCTG	594
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
DB	595	CAGCTTCGAAGCCTCACAGATCTTGAAGAGAGGGGTCTTGATCCACGCGGAACCCCCAG	654
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	180

Db	655	CTCTGCTACACGACACGATTTTGTGGAAGGACATCTCCACAAGAACACACGCTGGCT	714
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	715	CTCACACTATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG	774
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	775	GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTGTACAGAGCTGACGCGCACTGCTGT	834
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGluGlnCys	240
Db	835	GCGGTGGTGTGCTGCGCTGCAAGGGGCCACTGCCACCTGACTGCTGCCATGAGCAGTGT	894
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	895	GCTGCCGGCTGCACGGGGCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCAACCA	954
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	955	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACCTACAACACAGACACGTTTGAG	1014
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	1015	TCCATGCCCAATCCCGAGGGCCGGTATACATTGGGGCCGAGCTGTGTGACTGCTGTCCC	1074
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1075	TACAACCTACTTTCTACGGACGTGGGATCCTGCACCCCTGCTGCCCCCTGCACAAACCA	1134
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1135	GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA	1194
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1195	GTGTGCTATGGTCTGGGCATGGAGCACCTTGCAGAGGTGAGGGCAGTTTACCAGTGC	1254
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1255	ATCCAGGATTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC	1314
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe	400
Db	1315	TTTGATGGGAGCCACGCTCCAACACTGCCCGCTCCAGCCAGCAGCAGCTCCAAGTGT	1374
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1375	GAGACTCTGGAAGATACAGATTACCTATACATCTCAGCATGCGCGACAGCCTGCCT	1434
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1435	GACCTCAGCGTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACATGCGGCC	1494
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1495	TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGGCTGGCTCACTGAGGGAA	1554
QY	461	LeuGlySerClyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1555	CTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACTCTGCTTCGTGCACACGGTG	1614
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1615	CCCTGGGACCAGCTCTTTGGAAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1674
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1675	GAGGACGAGTGTGGGGCAGGGCCCTGGCTGCCACCACTGTCGCGCCCGAGGGCACTGC	1734
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1735	TGGGGTCCAGGGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCTTCCGGGGCCAGGAGTGC	1794
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1795	GTGGAGGAATCCGAGTACTGCAGGGGCTCCCCAGGAGTATGTGAATGCCAGGCACTGT	1854
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1855	TTGGCGTGCACCCCTGAGTGTACGCCCAAGATGGTCACTGACCTGTTTGGACCGGAG	1914
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1915	GCTGACCACTGTGGCCCTGTGCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC	1974
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1975	CCCAGGGTGTGAACCTGCCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2034
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2035	GGCGCATGCGACGCTTGGCCCATCAACTGCACCCACTCCTGTGTGACCTGGATGACAA	2094
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2095	GGCTGCCCGCCGAGCAGAGAGCCCTCTGACGTCCATCATCTCTCGGTGGTGGC	2154
QY	653	-----	653
Db	2155	ATTCTGCTGGTCTGTGGGTGGTGTCTTTGGGATCCTCATCAAGGACGGCAGCAG	2214
QY	653	-----	653
Db	2215	AAGATCCGGAAGTACACGATGCGGAGCTGCTGCAGAAACGGAGCTGTGGAGCCGCTG	2274
QY	653	-----	653
Db	2275	ACACCTAGCGGAGCGATGCCCAACAGCGCAGATCGGATCCTGAAAGACGCGGAGCTG	2334
QY	653	-----	653
Db	2335	AGGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGCTTACAAGGGCATCTGGATC	2394
QY	653	-----	653
Db	2395	CCTGATGGGGAGAAATGTAAAAATTCAGTGGCCATCAAAAGTGTGTAGGGAAAAACACATCC	2454
QY	653	-----	653
Db	2455	CCCAAGCCACAAGAAATCTTAGACGAAGCATACGTATGGCTGGTGGCTCCCCA	2514
QY	653	-----	653
Db	2515	TATGTCTCCGCCCTTCTGGGCATCTGCCCTGCATCCACGGTGCAGCTGGTGACACAGCTT	2574
QY	653	-----	653
Db	2575	ATGCCCTATGCTGCTCTTAGACCATGTCCGGGAAAAACCGGAGCGCTGGGCTCCCA	2634
QY	653	-----	653
Db	2635	GACCTGCTGAAGTGGTGTATGCAGATTGCCAAGGGGATGAGTACCTGGAGGATGTGGCG	2694
QY	653	-----	653
Db	2695	CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCACCATGTCAAA	2754
QY	653	-----	653
Db	2755	ATTACAGACTTGGGGCTGGCTGGCTGTGGACATTGACGAGACAGATACCATGCAGAT	2814
QY	653	-----	653
Db	2815	GGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTTCTCCGCCCGCGGTTACC	2874

```
Qy 653 ----- 653
Db 2875 CACCAGAGTGTGTGGAGTTATGTTGACTGTGTGGAGCTGATGACATTTTGGGGCC 2934
Qy 653 ----- 653
Db 2935 AACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGTGGAAGGGGAGCGG 2994
Qy 653 ----- 653
Db 2995 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTTCAATGTTGGATG 3054
Qy 653 ----- 653
Db 3055 ATTGACTCTGAATGTCTGGCCCAAGATTCCGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC 3114
Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 3115 AGGAGCCCCAGCCGCTTTGTGGTCCATCCAGAAATGAGGACTTGGGCCCGCAGCTCCCTTG 3174
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3175 GACAGCACCTTCTACCCGCTCAGTCTGGAGGACCATGACATGGGGGACCTGGTGGATGCT 3234
Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3235 GAGGAGTATCTGGTACCCAGCAGGGCTTCTTGTCTCCAGACCTGCCCGGGGCGCTGGG 3294
Qy 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
Db 3295 GGCATGTCTCCACACAGCCAGCCGAGCTCATCTACAGGAGTGGCGGTGGGGAGCTGACA 3354
Qy 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3355 CTAGGGCTGAGCCCTCTGAAGAGGAGGCCCCAGGTCTCCACTGGCACCTCCGAAGG 3414
Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3415 GCTGGCTCCGATGATTTGATGTGACCTGGGAATGGGCGAGCCAAAGGGGCTGCAAGC 3474
Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3475 CTCCTCCACATGACCCAGCCCTCTACAGCGGTACAGTGGAGGCCCCACAGTACCCTG 3534
Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3535 CCTCTGAGACTGATGCTAGCTAGTGGCCCCCTGACCTGCAGCCCCCAGCTGATATGTG 3594
Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3595 AACCAGCCAGATGTTGGCCCCCAGCCCTTCGCCCGGAGAGGGCCCTCTGCCCTGTGCC 3654
Qy 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3655 CGACCTGCTGGTGGCACTCTGGAAAGGCCCAAGACTCTCTCCCGAGGGAAGTGGGGTC 3714
Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3715 GTCAACAGCGTTTGGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3774
Qy 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3775 GGAGGAGCTGCCCTCAGCCCCACCCCTCTCTGCCCTTACGCCCGCAGCTTCGACACCTC 3834
Qy 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3835 TATTACTGGGACCAAGGACCACCAAGAGCGGGGGCTCCACCCAGCACCTTCAAAGGACA 3894
Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3895 CCTACCGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG 3939
```

RESULT 9

```
AAZ31071
ID AAZ31071 standard; DNA; 4473 BP.
AC AAZ31071;
DT 17-JAN-2000 (first entry)
DE HER-2 nucleic acid sequence.
KW HER-2; c-neu; ErbB2; transmembrane receptor; tyrosine kinase activity;
KW epidermal growth factor receptor; EGFR; HER-1; cancer; breast cancer;
KW ovarian cancer; gastric cancer; antisense oligonucleotide; expression;
KW hyperproliferative disease; ss.
XX Homo sapiens.
XX WO9948906-A1.
XX 30-SEP-1999.
XX 25-MAR-1999; 99WO-US06492.
XX 26-MAR-1998; 98US-0048804.
XX (ISIS-) ISIS PHARM INC.
XX (PENN-) PENN STATE RES FOUND.
XX Bennett CF, Lipton A, Witters LM;
XX WPI; 1999-610749/52.
XX New antisense sequences used to treat hyperproliferative conditions,
XX especially cancer -
XX Examples; Page 38-39; 44pp; English.
XX This is the human HER-2 polynucleotide sequence. The HER-2 gene also
XX called c-neu and ErbB2, encodes a transmembrane receptor, with tyrosine
XX kinase activity. HER-2 is related to the epidermal growth factor
XX receptor (EGFR or HER-1). Aberrant HER-2 expression is present in a wide
XX number of cancers, especially breast, ovarian and gastric cancers. This
XX sequence is used in the invention to design 12-25 nucleotide
XX oligonucleotides that decrease the expression of human HER-2. The
XX oligonucleotides of the invention (AAZ31067-231070) can also be used for
XX modulating the expression of human epidermal growth factor receptor. The
XX oligonucleotides are used to treat diseases or conditions associated with
XX HER-2, particularly hyperproliferative diseases such as cancer.
XX SQ Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other;
Alignment Scores:
Pred. No.: 7,31e-227 Length: 4473
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 20 Gaps: 1
SEQ6 (1-919) x AAZ31071 (1-4473)
Qy 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGAGCTGGGGCTGTGCTGGCGTGGGGCTCTCTCTCGCCCTTGTGCCCGGAGCC 234
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCCCAAGTGTGCACCGGCACACATGAAGCTGGCGCTCCCTGCAGTCCGAG 294
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAAACCTG 354
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
```


D	b	1435	GACCTCAGCGTCTTCACAAACCCTGC AAGTAATCCGGGAGCAAGATTCTGCACAATGGCGCC	1494
Q	y	441	Tyr-SerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
D	b	1495	TACTCGTGACCTGTCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCCTCACATGAGGGAA	1554
Q	y	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
D	b	1555	CTGGCAGTGGACTGGCCCTCATCCACCATAAACCCACCCTCTGCTCTGCTGCACACGGTG	1614
Q	y	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
D	b	1615	CCCTGGACCAGCTCTTTCCGAACCCGCCAAGCTCTGTCCACACTGCCAACCCGCCCA	1674
Q	y	501	GluAspGlnCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
D	b	1675	GAGGACAGTGTGTGGCGAGGGCTGGCCTGCCACCAGCTGTGCGCCCCGAGGCACTGC	1734
Q	y	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
D	b	1735	TGGSGTCCAGGGCCCACCCAGTGTGTCACTGAGCCAGTCTCTTCGGGCCAGGAGTGC	1794
Q	y	541	ValGluClnCysArgValLeuGlnGlyLeuProArgGluTyrrValasnAlaArgHisCys	560
D	b	1795	GTGGAGAAATGCCAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACGT	1854
Q	y	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
D	b	1855	TTGCGGTGCCACCCCTGAGTGTACGCCCCAGAATGGCTCAGTAGCCTGTTTGGACCGGAG	1914
Q	y	581	AlaAspGlnCysValAlaCysAlaHisTyrrLysAspProProPheCysValAlaArgCys	600
D	b	1915	GCTGACCACTGTGTGGCTGTGCCCATATAAGGACCTCCTCTCTGCGTGGCCCGCTGC	1974
Q	y	601	ProSerGlyValLysProAspLeuSerTyrrMetProIleTrpLysPheProAspGluGlu	620
D	b	1975	CCCAGCGTGTAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2034
Q	y	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValaspLeuAspAspLys	640
D	b	2035	GGCGCATGCCAGCGCTTGCCCCATCACTGCCACCCTCTCTGTGTGGACCTGGATGACAG	2094
Q	y	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer-----	653
D	b	2095	GGTGCCCCCGCAGAGAGAGAGCCCTCTCAGCTCCATCATCTCTGCGGTGGTTGGC	2154
Q	y	653	-----	653
D	b	2155	ATTCTGCTGCTGTGTCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGACAG	2214
Q	y	653	-----	653
D	b	2215	AAGATCGGAAGTACAGATGCGGAGACTGCTGCAGGAAACGGAGTGCTGGAGCCGCTG	2274
Q	y	653	-----	653
D	b	2275	ACACTAGCGGAGGGATGCCCAACCAAGCGCCACATGCGGATCCTGAAGAGACGGAGCTG	2334
Q	y	653	-----	653
D	b	2335	AGGAAGTGAAGTGTCTGGATCTGGGCTTTTGGCACAGTCTACAAGGSCATCTGGATC	2394
Q	y	653	-----	653
D	b	2395	CCTGATGGGAGAAATGTGAAAATTCAGTGGCCATCAAAGTGTGTGAGGGAAAAACATCC	2454
Q	y	653	-----	653
D	b	2455	CCCAAGCCCAACAAAGAAATCTTAGAGGARGATAGTGTATGGCTGGTGTGGGCTCCCCA	2514
Q	y	653	-----	653
D	b	2515	TATGCTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2574

Qy 653 ----- 653
 Db 2575 ATGCCCTATGGCTGCCCTTAGACCATGTCCGGGAAACCCCGGACGCCCTGGGGTCCAG 2634
 Qy 653 ----- 653
 Db 2635 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2694
 Qy 653 ----- 653
 Db 2695 CTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGTCAAGAGTCCCAACCATGTCAAA 2754
 Qy 653 ----- 653
 Db 2755 ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGACAGAGTACCATGCAGAT 2814
 Qy 653 ----- 653
 Db 2815 GGGGCAAGGTGCCCATCAAGTGTGATGGCGCTGGAGTCCATTCTCCGCCGCGGTTCACC 2874
 Qy 653 ----- 653
 Db 2875 CACCAGAGTGATGTGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2934
 Qy 653 ----- 653
 Db 2935 AAACCTTACGATGGATCCAGCCCGGGAGATCCCTGACCTGTGGAAAGGGGACGG 2994
 Qy 653 ----- 653
 Db 2995 CTGCCCCAGCCCCCATCTGCACCATGTATGTCTACATGATCATGTGTCAAATGTGGATG 3054
 Qy 653 ----- 653
 Db 3055 ATTGACTCTGAATGTCGGCCAGATTCGGGAGTTGGTGTCTGAATTCCTCCCGCATGGCC 3114
 Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 Db 3115 AGGAGCCCCAGCGCTTTGGTTCATCCAGATGAGGACTTGGCCCGCAGCTCCCTTG 3174
 Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 Db 3175 GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGTGGATGCT 3234
 Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
 Db 3235 GAGGAGTATCTGGTACCCCGACAGCGGCTTCTGTCTCCAGACCTGCCCCGGCGCTGGG 3294
 Qy 705 GlyMetValHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
 Db 3295 GGCATGTCACACAGCCAGCCGCTCATCTACAGGAGTGGCGTGGGGACCTGACA 3354
 Qy 725 LeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 Db 3355 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCGCCAGGTCTCCACTGGCACCTCCGAAGGG 3414
 Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyGlyLeuGlnSer 764
 Db 3415 GCTGGCTCCCATGATGATTTGATGTGACCTGGGAATGGGGACGCCAAGGGGCTGCAAGC 3474
 Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 Db 3475 CTCCCCACACATGACCCACCCCTCTACACGGGTACAGTGAGGACCCACAGTACCCCTG 3534
 Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 Db 3535 CCCTCTGAGACTGATGCTAGCTAGCTGCCCCCTGACCTGCAGCCCCCAGCCTGATATGTG 3594
 Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
 Db 3595 AACCAGCCAGATCTGGCCCCAGCCCCCTTCGCCCGCAGAGGGGCCCTCTGCCCTGTGCC 3654

825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
 Db 3655 CGACCTGCTGGTGGCACTCTGGAAGGCCCAAGACTCTCTCCCGAGGAGATGGGGTC 3714
 Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
 Db 3715 GTCAAGACGCTTTTGGCTTTGGGGTGGCGTGAGAACCCCGAGTACTTGCACACCCAG 3774
 Qy 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3775 GGAGGAGCTGCCCTCAGCCCCACCTCTCTGCTTCCAGCCAGCTTCGACAACTC 3834
 Qy 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
 Db 3835 TATTACTGGACACGAGGACCCACAGAGCGGGGGCTCCACCAGCACCTTCAAAGGACA 3894
 Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3895 CCTACGCGAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG 3939

RESULT 10
 ABZ34969
 ID ABZ34969 standard; cDNA; 4473 BP.
 AC ABZ34969;
 DT 05-FEB-2003 (first entry)
 XX Human gene expression profile polynucleotide SEQ ID NO 81.
 DE Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
 KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
 KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
 KW gene expression; gene; ss.
 XX Homo sapiens.
 XX WO200274979-A2.
 XX 26-SEP-2002.
 XX 20-MAR-2002; 2002WO-US08456.
 XX 20-MAR-2001; 2001US-276947P.
 XX (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
 XX Wan J, Wang Y;
 XX WPI; 2002-740862/80.
 XX New gene expression profile generated from primary, endothelial,
 PT epithelial, and muscle cell types, useful for identifying disease
 PT pathologies involving alterations of gene expression, e.g. cancer
 XX
 PS Claim 3; Page 274-276; 850pp; English.
 XX
 CC The invention relates to a gene expression profile comprising one or more
 CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
 CC is a coronary artery endothelium, umbilical artery or vein endothelium,
 CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
 CC endothelium, myometrium microvascular endothelium, keratinocyte
 CC epithelium, bronchial epithelium, mammary epithelium, prostate
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
 CC small airway epithelium, renal epithelium, umbilical artery smooth
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
 CC osteoblasts or prostate stromal cell. The gene expression profile is used
 CC for determining the level of RNA expression for a sample, determining the
 CC phenotype of a cell and distinguishing cell types. The gene or a protein
 CC expression profile is useful in identifying disease pathologies

CC involving alterations of gene expression. The assessment of expression
 CC profiles may provide meaningful information with respect to tumour type
 CC and stage, treatment methods, and prognosis. The gene or protein
 CC expression profile may also be used for creating microarrays. The
 CC microarray is useful for genetic and physical mapping of genomes, DNA
 CC sequencing, genetic or medical diagnosis, genotyping of organisms,
 CC confirming cell or tissue identifications and in identifying promising
 CC antibiotics, antiviral or antifungal agents.

XX
 SQ Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other;

Alignment Scores:
 Pred. No.: 7,31e-227 Length: 4473
 Score: 4900.00 Matches: 919
 Percent Similarity: 73.23% Conservative: 0
 Best Local Similarity: 73.23% Mismatches: 0
 Query Match: 96.49% Indels: 336
 DB: 24 Gaps: 1

SE06 (1-919) x AB234969 (1-4473)

QY 1 MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 DB 175 ATGGAGCTGGCGGCTTGGCGCGCTGGGGCTCTCTCGCCCTCTTGGCCCCCGGAGCC 234
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 235 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 294
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 DB 295 ACCCACTGGACATGCTCGCCACCTCTACCAAGGCTGCCAGGTGTGCGAGGAACCTG 354
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
 DB 355 GAACCTACCTACTGCTCCACCAATGCCAGCTGTCTTCTCCAGGATATCCAGAGGTG 414
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 415 CAGGGCTAGTGTCTCATGCTCACAAACCAAGTGAAGGAGGCTCCCACTGCAGAGGTGCGG 474
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 475 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGTAGACAATGGA 534
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB 535 GACCCGCTGAACAATACACCCCTGTACAGGGGCTCCCAAGAGGCTGCGGGAGCTG 594
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 DB 595 CAGCTTCGAAGCTCACAGAGATCTTGAAGGAGGGTCTTATCCAGCGGAACCCCGAG 654
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
 DB 655 CTTCTGCTACCGAGACACATTTTGTGGAAGGACATCTTCCACAAGAACAACAGTGGCT 714
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 715 CTCACACTGATAGACACACACACCCCTCTCTGGGCTGCCACCCCTGTCTCCGATGTATAG 774
 QY 201 GlySerArgCysTrpGlyLeuSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 775 GGCTCCCGCTGCTGGGAGAGATCTGTAGGATTGTGAGAGCTGTACGCGCTGTCTGTCT 834
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCys 240
 DB 835 GCGGTGGCTGTGTCGCGCTGCAAGGGGCGCACTGCCCACTGACTGCTGCTGCTGCTGCT 894
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 895 GCTGCGGCTGCACGGGCCCCAAGCACTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954

QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 955 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACAGACAGCTTTGAG 1014
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 1015 TCCATGCCCAATCCCGAGGGCGGTATACATTCCGGCGCAGCTGTGTACTCCCTGTCC 1074
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 1075 TACAACCTCTTCTACGGAGCTGGATCTCTCACCTCTGTCTGCCCCCTGCACACCAAA 1134
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 1135 GAGGTGCACAGCAGAGGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCGA 1194
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1195 CTGTCTATGTGTCTGGCATGGAGCACTTGCAGAGAGTGTGAGGCGAGTTACCACTGCCAAT 1254
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1255 ATCCAGAGTGTGTCTGGCTGCAAGAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC 1314
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB 1315 TTTGATGGGAGCCAGCCCTCCAAACCTGCCCGCTCCAGCCAGCAGCTCCCAAGTGT 1374
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 DB 1375 CAGACTCTGGAAGAGATCACAGTTTACCTATATCATCTCAGCATGTGCGCGCAGACGCTGCCT 1434
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1435 GACCTCAGCGCTCTCCAAACCTGCAAGTAATCCGGGAGCAATCTCTGCACAAATGGCGCC 1494
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1495 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGAA 1554
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisLysAsnThrHisLeuCysPheValHisThrVal 480
 DB 1555 CTGGGCATGTGACTGGCCCTCATCCACCATAACACCCCTCTGTCTGTGTCACAGGTG 1614
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1615 CCCTGGGAGCAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCA 1674
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1675 GAGGACGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGCTGCCCGCGAGGGCACTGC 1734
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1735 TGGGTTCAGGGGCCACCCAGTGTCTCAACTGCAGCCAGTTCCTTCTGGGGCCAGAGTGC 1794
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 DB 1795 GTGGAGGAATGCCAGTACTGCGAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1854
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 DB 1855 TTGCGCTGCCACCTGAGTGTGAGCCCAAGATGGCTCAGTCACTGTTTGTGGACCGGAG 1914
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 DB 1915 GCTGACCACTGTGTGGCTGTGCCACTATAAGGACCTCTCCCTTCTGCGTGGCGCGCTGC 1974
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 DB 1975 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCACTCTGGAAGTTTCCAGATGAGGAG 2034
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640

```
Db      2035  |||||GGCGATGCCAGCCTTGCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAG 2094
Qy      641  GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
Db      2095  GGTGCCCGCGGACAGAGACCGAGCCCTCTGACGTCATCATCTCTGCGGTGGTTGGC 2154
Qy      653  ----- 653
Db      2155  ATTCTGCTGGTGGTCTTGCGGTGCTTTGGGGTGCTTTGGGATCCTCATCAAGGACGGCAGCAG 2214
Qy      653  ----- 653
Db      2215  AAGATCCGGAAGTACACGATCGGAGACTGCTGCAGGAACGGAGCTGGTGGAGCCGCTG 2274
Qy      653  ----- 653
Db      2275  ACACCTAGCGGAGCGATGCCCAACACGAGCGGAGATGCGGATCGCTGAAAGACGAGGAGCTG 2334
Qy      653  ----- 653
Db      2335  AGGAAGGTGAAGGTCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2394
Qy      653  ----- 653
Db      2395  CCTGATGGGGAGAAATGTGAAATTCACAGTGGCCATCAAAGTGTGGAGGGAACACATCC 2454
Qy      653  ----- 653
Db      2455  CCCAAGCCACAAAGAAATCTTAGCAAGACGATACGTGATGGGTGGTGGCTGCCCA 2514
Qy      653  ----- 653
Db      2515  TATGTCCTCCGCTTCTGGGCATCTGCTGACATCCACGGTGACCTGGTGACACAGCTT 2574
Qy      653  ----- 653
Db      2575  ATGCCCTATGCTGCTCTTAGACCATGTCGGGAAACCGGAGCGCTGGGTCCCG 2634
Qy      653  ----- 653
Db      2635  GACCTGCTGAATGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2694
Qy      653  ----- 653
Db      2695  CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGTGTCAGAGTCCCAACCATGTCAA 2754
Qy      653  ----- 653
Db      2755  ATTACAGACTTCCGGCTGGCTCGGCTGTGGACATTTGACGAGACAGATGACCATGCAGAT 2814
Qy      653  ----- 653
Db      2815  GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATCTCCGCGCGGGTTTACC 2874
Qy      653  ----- 653
Db      2875  CACCAGACTGATGTGTGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGCCC 2934
Qy      653  ----- 653
Db      2935  AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGGAGCGG 2994
Qy      653  ----- 653
Db      2995  CTGCCCCAGCCCCCATCTGCACCATGATGCTACATGATCATGGTCAAAATGTTGGATG 3054
Qy      653  ----- 653
Db      3055  ATTGACTCTGAATGTGCGCCAAAGATTCCGGGAGTTGGTGTCTGAATCTCCCGCATGGCC 3114
Qy      654  -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
|||||
```

```
Db      3115  AGGACCCCCAGCGCTTTGTGTCTATCCAGATGAGGACTTTGGGCCCCAGCCAGTCCCTTG 3174
Qy      665  AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
|||||
Db      3175  GACAGCACCCTTCTACCGCTCACTGCTGGAGACGATGACATGGGGAGACTTGTGTGATGCT 3234
Qy      685  GluGlyTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
|||||
Db      3235  GAGGAGTATCTGGTACCCACGACGAGGCTTCTCTGTCTCCAGACCTTCCCGCGCGCTGG 3294
Qy      705  GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
|||||
Db      3295  GGCATGTTGCCACACAGGCACGACGCTCATCTACAGAGTGGGGTGGGACCTGACA 3354
Qy      725  LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGlyGly 744
|||||
Db      3355  CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCTCCAGGTCTCCACTGGCACCTTCCGAAGG 3414
Qy      745  AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
|||||
Db      3415  GCTGGCTCCGATGTATTTGATGTCGACCTGGGAATGGGCGAGCCCAAGGGCTGCAAGC 3474
Qy      765  LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
|||||
Db      3475  CTCCCAACACATGACCCAGCCCTCTACAGCGGTACAGTGAGGACCCACAGTACCCCTG 3534
Qy      785  ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
|||||
Db      3535  CCCCTGAGACTGATGGCTAGTGTGCCCTGACCTGACCTGAGCCCTCCAGCTGGAATATGTG 3594
Qy      805  AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
|||||
Db      3595  AACCCAGCAGATGTCGCGCCACGCCCCCTTCGCCCGCAGAGGGCCCTCTGCCTGCTGCC 3654
Qy      825  ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db      3655  CGACTGTGTGTGCACCTCTGGAAGGCCCAAGACTCTCTCCCCAGGGAAGAAATGGGTC 3714
Qy      845  ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db      3715  GTCAAGACGTTTTTGGCTTTGGGGTGCCGTGGAGAACCCTGAGTACTTGACACCCAG 3774
Qy      865  GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
|||||
Db      3775  GGAGGAGTGCCTCAGCCCCACCTCTCTCTCTCTGCTTTCAGCCCGCAGCTTCGACACCTC 3834
Qy      885  TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
|||||
Db      3835  TATTACTGGGACAGGACCCACAGAGCGGGGGCTCCACCAGCACCTTCAAGAGGACA 3894
Qy      905  ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
|||||
Db      3895  CCTACGCGAGAGAACCCAGAGTACCTGGGTCTGGACGTTGCCAGTG 3939

RESULT 11
AAD38904
ID  AAD38904 standard; DNA; 4473 BP.
AC  AAD38904;
XX
DT  23-SEP-2002 (first entry)
DE  Human Her-2 DNA.
XX
KW  Human; Her-2; epidermal growth factor receptor 2; infection; cancer;
KW  hyperproliferative disorder; prophylaxis; inflammation; antisense;
KW  tumour; gene therapy; phosphorothioate backbone; gene; ss.
OS  Homo sapiens.
FH  Key
FT  CDS 175..3942
FT  /*tag= a
```


QY	501	GlucpGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	Db	2755	ATTACAGACTTCGGGCTGGCTGGCTGGACATTGACGACAGAGTACCATGCAGAT	2814
Db	1675	GAGGACGAGTGTGTGGGAGGGCTGGCTGCCACCAAGCTGTGGCCGAGGCACTGC	1734	QY	653	-----	653
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540	Db	2815	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGGGGGTTCACC	2874
Db	1735	TGGGTCCAGGGCCCAACCACTGTCTAACTGCAAGCCAGTTCTTCGGGGCCAGAGTGC	1794	QY	653	-----	653
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560	Db	2875	CACCAGAGTGATGTCTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACATTTTGGGGCC	2934
Db	1795	GTGGAGGAATGCCGAGTACTCAGGGGCTCCCAAGGAGTATGTGAATGCCAGCACTGT	1854	QY	653	-----	653
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580	Db	2935	AAACCTTACGATGGGATCCCAGCCGGGAGATCCCTGACCTGCTGGAAGGGGGAGCGG	2994
Db	1855	TTGCGCTGCCACCTTGAGTGTGAGCCAGCCAGCAATGGCTCAGTCAGCTGTTTGGACCGGAG	1914	QY	653	-----	653
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600	Db	2995	CTGCCCCAGCCCCCATCTGCACCAATTGATGTCTACATGATCATGGTCAAAATGTTGGATG	3054
Db	1915	GCTGACCAAGTGTGGCCCTGTGCCCACTATAGGACCTCTCTCTGGTGGCCGCTGC	1974	QY	653	-----	653
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620	Db	3055	ATTGACTCTGAATGTGCGCCCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCGCATGGCC	3114
Db	1975	CCCAGCGGTGAAACCTGACCTCTCTTACATGCCATCTGGAAGTTTCCAGATGAGGAG	2034	QY	654	-----	654
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640	Db	3115	AGGACCCCCAGCGCTTGTGGTCATCCAGAAATGAGGACTTGGGCCACAGCAGTCCCTTG	3174
Db	2035	GGCGCATGCCAGCTTGCCCCATCAACTGCAACCCACTCTCTGTGGACCTGGATGACAAG	2094	QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653	Db	3175	GACAGCACTTCTACCGCTCACTGCTGGAGACGATGACATGGGGGACCTGGTGGATGCT	3234
Db	2095	GGCTGCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGGCGTGGTTGGC	2154	QY	685	GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly	704
QY	653	-----	653	Db	3235	GAGGAGTATCTGGTACCCACAGCGGCTTCTTCTGTCAGACCTGCCCGGGCGCTGGG	3294
Db	2155	ATTCTGCTGGTGGTCTTGGGGTGGTCTTTGGGATCTTCATCAAGCGACGACGAG	2214	QY	705	GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr	724
QY	653	-----	653	Db	3295	GGCATGTGTCCACCAACAGCACCGACCGAGCTCATCTACAGGAGTGGCGGTGGGACCTGACA	3354
Db	2215	AAGATCCGGGAAGTACAGGATCGGAGACTGCTGCGGAAAGGAGCTGGTGGAGCCGCTG	2274	QY	725	LeuGluLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
QY	653	-----	653	Db	3355	CTAGGGCTGGAGCCCTCTGAAGAGAGGGCCCCAGGTTCTCCACTGGCACCTCCGANGGG	3414
Db	2275	ACACCTAGCGAGCGATGCCCAACAGCGCGAGATGCGGATCTGTGAAGAGACGGAGCTG	2334	QY	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
QY	653	-----	653	Db	3415	GCTGGCTCCGATGTATTTGATGGTGACCTGGGAATGGGGAGGCAACAGGGCTGCAAGC	3474
Db	2335	AGGAAGTGAAGTGTGGATCTGGCCCTTTTGGCACAGTCTACAGGGGATCTGGATC	2394	QY	765	LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu	784
QY	653	-----	653	Db	3475	CTCCCCACATGACCCAGCCCTCTACAGCGGTACAGTAGAGGACCCACACAGTACCCCTG	3534
Db	2395	CCTGATGGGAGAAATGTGAATAATTCAGTGGCCATCAAAAGTGTGAGGGAAACACATCC	2454	QY	785	ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal	804
QY	653	-----	653	Db	3535	CCCTCTGAGACTGATGGCTACGTTGGCCCCCTGACCTGCAGCCCCCAGCCCTGGAATATGTG	3594
Db	2455	CCCAAGCCAAAGAANAATTTAGACGAAGCATACGTGATGGTGTGGGTCCCCCA	2514	QY	805	AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla	824
QY	653	-----	653	Db	3595	AACCAAGCAGATGTCGGCCCCAGCCCTTCGCCCCGAGAGGGCCCTCTGCCTGCTGCC	3654
Db	2515	TATGTCCTCCGCCCTTCTGGGATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2574	QY	825	ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal	844
QY	653	-----	653	Db	3655	CGACCTGTGTGGTCACTCTGGAAGGCCCAAGACTCTCTCCCGAGGAGAAATGGGTC	3714
Db	2575	ATGCCCTATGGCTGCCCTTATAGACCATGTCCGGGAAACCGCGACGCTGGGCTCCAG	2634	QY	845	ValLysAspValPheAlaPheGlyAlaValGluAsnProGluTyrLeuThrProGln	864
QY	653	-----	653	Db	3715	GTCAAGACGCTTTTGGCTTTGGGGTGGCTGGAGAACCCCGAGTACTTGACACCCAG	3774
Db	2635	GACCTGTGAAGTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGG	2694	QY	865	GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu	884
QY	653	-----	653	Db	3775	GGAGAGCTGCCCTCAGCCCCCACCCTCTCTGCTTCCAGCCAGCCTTCGACACCTC	3834
Db	2695	CTCCTACACAGGACTTGGCGCTCGGACGTGTGGTCAAGAGTCCCAACCATGTCAA	2754	QY	885	TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr	904
QY	653	-----	653				

Db 3835 TATTACTGGACAGGACCCACAGAGCGGGGGCTCCACCAGCACCTTCAAAGGGACA 3894
 QY 905 ProThraAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3895 CCTACGGCAGAAACCCAGAGTACCTGGGTCTGGACGTCCAGTG 3939

RESULT 12

AAA09455
 ID AAA09455 standard; DNA; 3768 BP.

XX AAA09455;

XX 10-AUG-2000 (first entry)

XX Human heregulin 2 (Her2) coding sequence.

XX Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
 KW self-protein; cell-associated peptide antigen; foreign epitope;
 KW cancer; breast cancer; prostate cancer; ss.

XX Homo sapiens.

XX W0200020027-A2.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-DK00525.

XX 05-OCT-1998; 98DK-0001261.

PR 20-OCT-1998; 98US-0105011.

XX (MEBI-) M & E BIORECH AS.

XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;

XX WPI; 2000-349917/30.

DR P-PSDB; AAY92620.

XX Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer

XX Claim 62; Page 187-193; 220pp; English.

XX The claims detail a method for inducing immune responses against weakly
 CC immunogenic cell-associated peptide antigens (PA) such as those
 CC associated with cancers (i.e. self-proteins), for example, human
 CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
 CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
 CC simultaneous presentation by antigen producing cells (APCs) of the
 CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
 CC group derived from the PA and/or at least 1 B-cell group derived from the
 CC cell-associated PA; and (2) at least 1 first T helper cell group which is
 CC foreign to the animal. Analogues of human PSM, human Her2 and
 CC human/murine FGF8b comprising a substantial part of all known and
 CC predicted CTL and B-cell epitopes of the respective PA and including at
 CC least one foreign T helper epitope are also claimed. The method is used
 CC to treat prostate, prostate/breast or breast cancer when the PA is human
 CC PSM, FGF8b and Her2, respectively.

XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:

Pred. No.: 1,48e-226 Length: 3768
 Score: 4892.00 Matches: 918
 Percent Similarity: 73.15% Conservative: 0
 Best Local Similarity: 73.15% Mismatches: 1
 Query Match: 96.34% Indels: 336
 DB: 21 Gaps: 1

SE06 (1-919) x AAA09455 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuProGlyAla 20

1	ATGGAGCTGGCGCCTTGTGGCGCTCCTCCTCGCCTCTTGCCTCCCGGAGCC	60	
21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40	QY
61	GCAGCACCACCAAGTGTGCGCGCACAGACATGAAGCTCGGCTCCCTCCCAAGTCCCGAG	120	Db
41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60	QY
121	ACCCACCTGGACATGCTCGCCACCTCTTACAGGGGCTGCAGGTGGTGCAGGAAACCTTG	180	Db
61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal	80	QY
181	GAACCTCACCTACCTGCCACCAATGCCAGCTGTCTCTCTCAGGATATCAGGAGGTG	240	Db
81	GlnGlyTyrValLeuIleAlaHisGlnValArgGlnValProLeuGlnArgLeuArg	100	QY
241	CAGGGCTAGTGTCTATCGCTCACAAACCAAGTGAGCAGGTCCCACTGCAGAGGCTGCGG	300	Db
101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120	QY
301	ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTAGCAATGGA	360	Db
121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140	QY
361	GACCCGCTGAACAATACCACCTGTGTACAGGGGCTCCCCAGGAGGCTGCGGGAGCTG	420	Db
141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgsnProGln	160	QY
421	CAGCTTCGAAGCTTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG	480	Db
161	LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla	180	QY
481	CTCTGCTACCAGACACAGATTTTGTGAAGAGACATCTTCCACAAGAACACACAGCTGGCT	540	Db
181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200	QY
541	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCGATGTGTAAG	600	Db
201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220	QY
601	GGCTCCCGCTGTGGGGAGAGAGATTCTGAGGATTGTGAGAGCTGACAGGCTGACGGCTGTCTGT	660	Db
221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240	QY
661	GCCGTGGCTGTGCCGCTGCAAGGGCCACCTGCCACCTGCTGCTGCCATGAGCAGTGT	720	Db
241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	QY
721	GCTGCCGGCTGCAGCGGCCCAAGCACTCTGACTGCTGCCCTGCCCTTCAACCCAC	780	Db
261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	QY
781	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTTCACCTACAAACACAGACAGCTTGTAG	840	Db
281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	QY
841	TCCATGCCCAATCCCGAGGGCCGATATCATTCGGGCCAGCTGTGTGACTGCTCTGCC	900	Db
301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	QY
901	TACAACTACCTTTCTACGGACCTGGGATCTCTGCACCTCTGCTGCCCTGCACCAACCAA	960	Db
321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	QY
961	GAGGTGCACAGCAGAGGATGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCCCGA	1020	Db
341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	QY
1021	GTGTGCTATGCTGTGGCATGGAGGACCTTGCAGAGAGGTGAGGCGAGTTACAGTGCCTAAT	1080	Db
361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGlyLeu	380	QY

Db	1081	ATCAGAGAGTTTGGCTGCGTGCACGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400
Db	1141	TTTGTATGGGACCCAGCGCTCCAACACTGCCCGCTCCAGCAGAGACGACGCTCCAAGTGT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTACGCTCTTCCAGAACCTGCAAGTAATCCGGGAGCAATTCGCACATGGCGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGGCTCACTGAGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGCAGTGGACTGGCGCTCATCCACATAACACCCACCTCTGCTTCGTGCACACGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAAGCTCTTTTCGGAACCCGCACCAAGCTCTGTCCACACTGCCAACCGGCCA	1500
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACGAGTGTGGGCGAGGGCTGGCTGCCACCAAGCTGTGGCGCCGAGGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGCCACCAGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1620
Qy	541	ValGluCysArgValLeuGlnGlyLeuProArgGluTrpValAsnAlaArgHisCys	560
Db	1621	GTGGAGAAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTGAGTGTACGCCCCAGCAATGGCTCAGTGACCTGTTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCAGTGTGTGGCTGTGCCACTATAAGGACCTCCCTCTCGTGGCGCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCGACGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCCAGCTTGCCCCATCACTGCACCCACTCTCTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCTGCTCGCGGTGGTGGC	1980
Qy	653	-----	653
Db	1981	ATTCTGCTGCTGGTGTGGGGTGGTCTTTTGGGATCTCATCAAGCGACGGCAGCAG	2040
Qy	653	-----	653
Db	2041	AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAACGGAGCTGGTGGAGCCGCTG	2100
Qy	653	-----	653
Db	2101	ACACCTAGCGGCGATGCCAACAGCGCAGATGCGGATCCTGAAAGAGACGGAGCTG	2160
Qy	653	-----	653
Db	2161	AGGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTTACAAGGGCATCTGGATC	2220
Qy	653	-----	653
Db	2221	CCTGATGGGAGAGATGTGAAAAATTCAGTGGCCATCAAAAGTGTTCAGGGAAAAACACATCC	2280
Qy	653	-----	653
Db	2281	CCCAAGCCAAACAAGAAATCTTAGAGGAAGCATACGTGATGGCTGTGTGGGCTCCCCA	2340
Qy	653	-----	653
Db	2341	TATGCTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400
Qy	653	-----	653
Db	2401	ATGCCCTATGGCTGCCCTCTTAGACCATGTCCGGGAACCCGCGACGCCCTGGGCTCCAG	2460
Qy	653	-----	653
Db	2461	GACCTGTGAATGGTGTATGCAGATTGCCAAGGGGATGAGCTACTGGAGGATGTGCGG	2520
Qy	653	-----	653
Db	2521	CTCGTACACAGGAGACTTGGCGCTCGGAACGTGCTGTCAGAGTCCCAACCATGTCAAA	2580
Qy	653	-----	653
Db	2581	ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTTCACGACAGACAGATACCATGCAGAT	2640
Qy	653	-----	653
Db	2641	GGGGCAAGTGGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGCTTCACC	2700
Qy	653	-----	653
Db	2701	CACCAGAGTATGTGTGGAGTTATGTGTGACTGTGTGGGAGTGATGACTTTTGGGGCC	2760
Qy	653	-----	653
Db	2761	AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGTGGAAGGGGGAGCGG	2820
Qy	653	-----	653
Db	2821	CTGCCCGACCCCCCATCTGCACCATTTGATGCTTACATGATCATGTGTCAAATGTTGGATG	2880
Qy	653	-----	653
Db	2881	ATTGACTCTGAATGTGCGGCCAAGATTCCGGGAGTTGTGTGTAATTCCTCCCGCATGGCC	2940
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	2941	AGGGACCCCGACGCTTTGTGGTCAATCCAGAATGAGGACTTGGGCCCGACGATCCCTTG	3000
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTCTACCGCTCAGCTGCTGGAGACGATGACATGGGGGACCTGGTGGATGCT	3060
Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGAGTATCTGGTACCCCGACGAGGCTTCTTGTCCAGACCTTCCCGGGCGCTGGG	3120
Qy	705	GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr	724
Db	3121	GGCATGGTCCACACAGGACCCGACGCTCATCTACAGGAGTGGCGTGGGACCTGCACA	3180
Qy	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	3181	CTAGGGCTGAGCCCTCTGAAGAGGAGGCCGCCAGGTCTCCACTGGCACCTCCGAGGG	3240
Qy	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
Db	3241	GCTGGCTCCGATGATTTTGTGTGACCTGGGAATGGGGGAGCCCAAGGGGCTGCAAGC	3300

QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 Db 3301 CTCCTGACATGACCCAGCCCTCTACAGCGTACAGTGGAGACCCACAGTACCCCG 3360
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 Db 3361 CCCTCTGAGACTGATGGCTACCTTGCCTCCCTGACCTGCAGCCCGCAGCTGAATATG 3420
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
 Db 3421 AACCGCCAGATGTCGGCCCGCCAGCCCTTGCCTCCCGAGAGGGCCCTCTGCTGCTGC 3480
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
 Db 3481 CGACCTGCTGGTGCCTACTGGAAGGGCCAGAGCTCTCTCCCGAGGGAAGTGGGTC 3540
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
 Db 3541 GTCAAGACGTTTTCCTTGGGCTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3600
 QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3601 GGAGGAGCTGCCCTCAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3660
 QY 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
 Db 3661 TATTACTGGACAGGACCCAGAGCGGGGCTCCACCCAGCACCTTCAAGGGACA 3720
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3721 CCTACGGCAGAACCCAGAGTACCTGGGTCTGGAGTCCGACGTCCAGTG 3765
 RESULT 13
 ABZ35744
 ID ABZ35744 standard; DNA; 3768 BP.
 XX
 AC ABZ35744;
 XX
 DT 07-FEB-2003 (first entry)
 XX
 DE Human ERBB2 polynucleotide SEQ ID NO 52.
 XX
 KW Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
 KW protozoicide; gene expression; antisense; tumour; infection; plasmodium;
 KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
 KW Hepatitis C virus; human papilloma virus; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN DE10100588-A1.
 XX
 PD 18-JUL-2002.
 XX
 PF 09-JAN-2001; 2001DE-1000588.
 XX
 PR 09-JAN-2001; 2001DE-1000588.
 XX
 PA (RIBO-) RIBOPHARMA AG.
 XX
 PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
 XX
 DR WPI; 2002-683450/74.
 XX
 XX Inhibiting expression of target genes, useful e.g. for treating tumors,
 PT by introducing into cells two double-stranded RNAs that are
 PT complementary to the target
 XX
 PS Claim 13; Page 38-39; 100pp; German.
 CC
 CC The invention relates to inhibiting expression of a target gene in a cell
 CC by introducing at least two oligonucleotides (dsRNAI and II), both
 CC with a double-stranded (ds) structure of at most 49 sequential nucleotide
 CC pairs. At least part of one strand (S1, S2) of the ds structures in each

CC of dsRNAI and II are complementary to regions in the target gene. The
 CC method uses antisense inhibition of gene expression using double stranded
 CC RNA inhibition (RNAi). The method is particularly used to treat tumours
 CC or infections, especially by Plasmodium or viruses/viroids (pathogenic on
 CC humans, animals or plants). The method provides more effective inhibition
 CC of expression than known methods using a single dsRNA, even at very low
 CC concentrations. When dsRNA has at least one unpaired nucleotide at the
 CC end, stability (and thus effective concentration in the cell) is
 CC improved and efficiency can be increased further by pretreating the cells
 CC with interferon. The present sequence is that of a target DNA of the
 XX invention.
 SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.48e-226 Length: 3768
 Score: 4892.00 Matches: 918
 Percent Similarity: 73.15% Conservative: 0
 Best Local Similarity: 73.15% Mismatches: 1
 Query Match: 96.34% Indels: 336
 DB: 24 Gaps: 1
 SEQ6 (1-919) x ABZ35744 (1-3768)
 QY 1 MetGluLeuAlaAlaLeuCysArgTyrGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 Db 1 ATGGAGCTGGCGGCTTGTGGCGCTGGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProLysProGlu 40
 Db 61 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGTCCGCTCTCTCTCTCTCTCTCT 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 121 ACCACCTGGAGATGCTCGCCACCTCTTACCAGGGCTGCCAGGTGGTGGAGGAACCTG 180
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 181 GAACCTCACCTACCTGCCACCAATGCCAGCTGTCTCTCTCTCTCTCTCTCTCTCTCT 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 241 CAGGGCTAGCTGCTCATCGCTCACACCAAGTGAAGGAGTCCCACTGCAGAGGTGGCG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 301 ATTGTGGAGGACCCAGCTCTTTGAGGACACTATGCCCTGGCCCTGTCTACAAATGGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 361 GACCCGCTGAACAATACACACCTGTCTACAGGGGCTCTCCAGAGGCGCTGGGGAGCTG 420
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 Db 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGTATCCAGGGGAACCCCG 480
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 Db 481 CTCTGCTACCGACACAGATTTTGTGAAGGACATCTTCCACAGAACCAACACAGCTGGCT 540
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 541 CTCACACTGATAGACACCAACCCGCTCTCGGGCTGCCACCTCTGTCTCTCCGATGTGTA 600
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 601 GGCTCCCGCTGCTGGGGAGAGAGTCTGAGGATTGTACAGAGCTGTACGCGCTACTGTCT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
 Db 661 GCCGTGGCTGTGCCCGCTGCAAGGGGCGACCTGCCACTGCTGCTGCTGCTGCTGCTG 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260

QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 Db 2941 AGGGACCCAGCGCTTGTGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTG 3000
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 Db 3001 GACAGACCTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGTGTGATGCT 3060
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
 Db 3061 GAGGAGTATCTGTACCCAGCAGGCTTCTGTGTCAGACCCCTGCCCCGGCGCTGGG 3120
 QY 705 GlyMetValHisHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
 Db 3121 GGCATGGTCCACACAGCAGCCAGCTCATCTACAGGAGTGGCGTGGGACCTGACA 3180
 QY 725 LeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 Db 3181 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCTCCAGTCTCCACCTCCGAAAGG 3240
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyGlyLeuGlnSer 764
 Db 3241 GCTGGCTCGATGATTTGATGTCACCTGGGAATGGGGGAGCCCAAGGGGCTGCAAGC 3300
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 Db 3301 CTCCCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAGGACCCACATACCCCTG 3360
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 Db 3361 CCTCTGAGACTGATGGCTACCTGTCCTCCCTGACCTGCAGCCCCCAGCTGAATATG 3420
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
 Db 3421 AACCAGCCAGATGTCGGCCCCAGCCCCCTTCGCCCGAGGAGGCCCTCTGCTGCTGCC 3480
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
 Db 3481 CGACCTGCTGTGTGCCACTCTGAAAGGGCCAGACTCTCTCCAGGGAAGAAATGGGTC 3540
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
 Db 3541 GTCAAGACGTTTTTGCCFTTGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCAG 3600
 QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3601 GGAGGAGCTGCCCTCAGCCCCACCTCTCTGCTTCCAGCCAGCTTCACACACCTC 3660
 QY 885 TyrTyrThrAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
 Db 3661 TATTACTGGGACAGCAGCCACAGAGCGGGGGCTCCACCCAGCACCTTCAAGGGACA 3720
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3721 CCTACGGCAGAACCCAGAGTACCTGGGTCTGGAGCTGCCAGT 3765
 RESULT 14
 ABX09987
 ID ABX09987 standard; DNA; 3768 BP.
 XX AC ABX09987;
 XX XX
 DT 23-JAN-2003 (first entry)
 XX DE Human ERBB2 DNA fragment SEQ ID 52.
 XX XX
 KW Oligoribonucleotide; interferon; oncogene; cytokine; id; developmental;
 KW KW prion; inhibition; human; ds.
 XX OS Homo sapiens.
 XX XX
 PN DE10100587-C1.

XX 21-NOV-2002.
 XX 09-JAN-2001; 2001DE-1000587.
 XX 09-JAN-2001; 2001DE-1000587.
 XX (RIBO-) RIBOPHARMA AG.
 XX Kreutzer R, Limmer S, Rost S, Hadwiger P;
 XX WPI; 2002-742209/81.
 XX Inhibiting expression of target genes, e.g. oncogenes, in cells, by
 PT introduction of complementary double-stranded oligoribonucleotide,
 PT after treating the cell with interferon
 XX Disclosure; Page 43-44; 98pp; German.
 CC This invention describes a novel method for inhibiting expression of a
 CC target gene by introducing into the cell that contains the target gene
 CC at least one oligoribonucleotide (dsRNA) that has a double-stranded
 CC (ds) structure of not more than 49 consecutive nucleotides (nt), where
 CC at least a segment of one strand of the ds structure is complementary
 CC with the target gene and the cells are treated with interferon before
 CC introduction of dsRNA. The method is used to inhibit expression of
 CC target genes, particularly oncogenes, cytokine genes, id (not defined)
 CC protein genes; developmental or prion genes, or genes expressed in
 CC pathogenic organisms (particularly plasmidia) or in viruses or viroids
 CC (pathogenic in humans, animals or plants). Treating the cells with
 CC interferon greatly increases the extent to which dsRNA can inhibit
 CC expression of the target genes, and the effect is even greater when dsRNA
 CC are modified to increase their stability. ABX09936-ABX10075 represent
 CC gene fragments used to illustrate the method of the invention.
 XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 1,48e-226 Length: 3768
 Score: 4892.00 Matches: 918
 Percent Similarity: 73.15% Conservatives: 0
 Best Local Similarity: 73.15% Mismatches: 1
 Query Match: 96.34% Indels: 336
 DB: 24 Gaps: 1
 SEQ6 (1-919) x ABX09987 (1-3768)
 QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
 Db 1 ATGGAGCTGGCGGCTTGTGGCGCTGGCGCTCTCTCGCCCTCTTGTCCCGCGAGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrGlyAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 61 GCGAGACCCAAAGTGTGCACCGGCACAGACATGAGTGGCGCTCCCTCCAGTCCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 121 ACCCACTGGACATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGGTGGAGGAAACCTG 180
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 181 GAACTCACCTACCTGCCCAATGCAAGCTGTCTCTCTCCAGGATATCCAGGAGGTG 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 241 CAGGGCTACGTGCTCATCGTCAACCAAGTGGAGGAGTCCCTACTGAGAGGTGCGG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCCCGCTGCTAGACAATGGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB: 140

Db	361	GACCCGCTGAACAATACACCCCTGTCTACAGGGGCTCTCCAGGAGGCGCTCGGGAGCTG	420	QY	501	GluAspGluCysValGlyGluAlaCysHisGlnLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160	Db	1501	GAGGACGAGTGTGTGGCGAGGGCCCTGCCCTGCCACAGCTGTGCGCCCGAGGGCACTGC	1560
Db	421	CACGCTCGAAGCCTCACAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCGAG	480	QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180	Db	1561	TGGGTGCCAGGGCCCAACCAGTGTGTCAACTGCAGCCAGTTCTCTCGGGCCAGAGATGC	1620
Db	481	CTCTGTACAGACACGATTTTGTGAAGAGACATCTTCCCAAGAACCAACAGCTGGCT	540	QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200	Db	1621	GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGATGCCAGGCACGTGT	1680
Db	541	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG	600	QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220	Db	1681	TTGCGCTGCCACCCCTGAGTGTGAGCCCAAGATGGCTCAGTGACCTGTTTGGACCGGAG	1740
Db	601	GGCTCCCGTGTGGGAGAGATGTTCTGAGGATTTCTCAGAGCCTGACGCGCACTGTCTGT	660	QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys	600
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys	240	Db	1741	GCTGACCAAGTGTGTGGCTGTGCCACATATAAGGACCCCTCCCTTCTGCTGGGCCGCTGC	1800
Db	661	GCCGGTGGGTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGTCGTCATGAGCAGTGT	720	QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	Db	1801	CCCAGGGTGTGAACCTGACCTCTCTACATGCCATCTCCAGTCTCGAAGTTCCAGATGAGAG	1860
Db	721	GCTGCCGCTGCACGGGCCCCAAGCACTCTGACTGCCCTGGCTGCCCTCCACTTCAACCAC	780	QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	Db	1861	GGCGCATGCCAGCCCTGCCCATCAACTCAACCCACCTCCTGTGTGGACCTGGATGACAAG	1920
Db	781	AGTGGCATCTGTGAGCTGCACCTGCCACGCCCTGGTCACTACAACACAGACACCTTTGAG	840	QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer-----	653
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	Db	1921	GGCTGCCCGCGCAGAGAGAGCCAGCCCTCTGAGCTCCATCGTCTCTGCGGTGGTGGC	1980
Db	841	TCCATGCCCAATCCCGAGGGCGGTATACATTTCGGCGCAGCTGTGTGACTGCTGCTCCC	900	QY	653	-----	653
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	Db	1981	ATTCTGCTGCTCGTGGTCTTGGGGTGGTCTTTTGGGATCTCTATCAAGCAGCGCAGCAG	2040
Db	901	TACAATACCTTTCTACGGACGTGGATCTCTGACCCCTCGTCTGCCCTGCACAACCAA	960	QY	653	-----	653
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	Db	2041	AAGATCCGGAAGTACACGATGCGGAGACTGTGTCAGGAAACGAGCTGGTGGAGCCGCTG	2100
Db	961	GAGGTACAGCAGAGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA	1020	QY	653	-----	653
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	Db	2101	ACACCTAGCGGAGCGATGCCCAACCAAGCGCAGATGCGGATCTCTGAAAGAGACGAGCTG	2160
Db	1021	GTGTGTATGGTCTGGCATGAGCATTGCGAGAGGTGAGGCGAGTTTACAGTGCCAAT	1080	QY	653	-----	653
QY	361	IleGlnGluPheAlaCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	Db	2161	AGGAAGGTGAAGGTGCTTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220
Db	1081	ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCCCTGGCATTTCTGCCGGAGAGC	1140	QY	653	-----	653
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400	Db	2221	CCTGATGGGAGAAATGTGAAATATCCAGTGGCCATCAAGTGTGTGAGGGAACACACATCC	2280
Db	1141	TTTGATGGGAGCCAGCCTCCAACTATACATCTCAGCATGCGCGCAGAGAGCCTCCAAGTGT	1200	QY	653	-----	653
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420	Db	2281	CCCAAGGCAACAAGAAATCTTAGACGAAGCATACGTTAGTGGTGTGGGTCTCCCA	2340
Db	1201	GAGACTCTGGAAGAGATCACAGTTTACCTATACATCTCAGCATGCGCGCAGAGCCTGCCT	1260	QY	653	-----	653
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	Db	2341	TATGTCTCCGCTTTCTGGGCATCTGCCCTGACATCCACGCTGCAGCTGGTGACACAGCTT	2400
Db	1261	GACCTCAGCGTCTTCAGAACCTGCAAGTAAATCCGGGACGAATCTGCACAAATGGCGCC	1320	QY	653	-----	653
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460	Db	2401	ATGCCCTATGGCTGCTCTTAGACCATGTCCTCGGGAAACCGCGGACGCCCTGGGCTCCCAG	2460
Db	1321	TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGGCTGCGCTACTGAGGGAA	1380	QY	653	-----	653
QY	461	LeuGlySerGlyLeuAlaIleHisHisAsnThrHisLeuCysPheValHisThrVal	480	Db	2461	GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2520
Db	1381	CTGGCAGTGGACTGGCCCTCATCCACCATACACCCACCTCTGCTTCGTGCACACGGTG	1440	QY	653	-----	653
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	Db	2521	CTCGTACACAGGAGT	

QY 653 ----- 653
 Db 2581 ATTACAGACTTCGGCTGGCTGGCTGGACATTGACGAGACAGATACCATGCAGAT 2640
 QY 653 ----- 653
 Db 2641 GGGGGCAAGGTGCCATCAAGTGGATGGGCTGGAGTCCATTCTCCGCGGGGTTTCACC 2700
 QY 653 ----- 653
 Db 2701 CACCAGATGATGTGGAGATTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2760
 QY 653 ----- 653
 Db 2761 AAACCTTACGATGGGATCCAGCCGGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2820
 QY 653 ----- 653
 Db 2821 CTGCCCCAGCCCCCATCTGCACCAATTGATGTCTACATGATCATGTCAAATGTTGGATG 2880
 QY 653 ----- 653
 Db 2881 ATTGACTCTGAATGTCCGCCAAGATTCCGGGAGTTGGTGTCTGAATTCCTCCCGCATGGCC 2940
 QY 654 -----
 Db 2941 AGGGACCCCGAGCGCTTTGTGGTTCATCCAGAATGAGGACTTGGCCCGACCCAGTCCCTTG 3000
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 Db 3001 GACAGACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGTGGATGCT 3060
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
 Db 3061 GAGGAGTACTGTACCCAGCAGCGGCTTCTCTGTCAGACCTGCCCGGCGCTGGG 3120
 QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
 Db 3121 GGCATGGTCCACCACAGCCAGCCAGCTCATCTACCAGGAGTGGCGTGGGACCTGACA 3180
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 Db 3181 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCGCCAGGTCTCCACTGGCACCTCCGAAGG 3240
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
 Db 3241 GCTGGCTCGATGATTTGATGTTGATGCTGGGATGGGCGACCCAGGGGCTGCANAGC 3300
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 Db 3301 CTCCCCACACATGACCCCGCCCTCTACAGCGGTACAGTGAGGACCCACAGTACCCCTG 3360
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 Db 3361 CCCTCTGACGATGATGGTACGTTGCCCCCTGACCTGACGAGCCCGCCAGCTGAATATG 3420
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
 Db 3421 AACCACCCAGATGTTGGCCCCCAGCCCCCTTGCCTCCGAGAGGCCCTCTGCTGCTGCC 3480
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
 Db 3481 CGACCTGCTGGTCCACTCTGGAAGGGCCAGACTCTCTCCCGAGGGAAGATGGGCTC 3540
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
 Db 3541 GTCAAAAGAGTCTTTTGCCTTGGGGTGGCTGGAGAACCCCGAGTACATTCACACCCAG 3600
 QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3601 GGAGGAGTGCCTCCTCAGCCCCCACCCTCTCTGCTCTCAGCCCTTCAGCCCTTCGACACCTC 3660
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904

Db 3661 TATTACTGGACCCAGGACCCAGAGCGGGGCTCCACCCAGACCTTCAAAGGACA 3720
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3721 CCTACGGCAGAACCCAGAGTACCTGGTCTGGAGTCCAGTG 3765
 RESULT 15
 AAD43935
 ID AAD43935 standard; cDNA; 3768 BP.
 AC AAD43935;
 DT 13-DEC-2002 (first entry)
 DE Human HER-2 cDNA.
 XX
 KW Transgenic animal; transgenic; mammary gland cell; HER2; tumour;
 KW cancer; therapy; apoptosis; cytostatic; human; gene; ss.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 1..3768
 FT /*tag= a
 FT /product= "Human HER2 protein"
 XX
 PN US2002035736-A1.
 XX 21-MAR-2002.
 XX 16-MAR-2001; 2001US-0811115.
 XX 16-MAR-2000; 2000US-189844P.
 XX (ERIC/) ERICKSON S.
 PA (KING/) KING K.
 PA (SCHW/) SCHWALL R.
 XX Erickson S, King K, Schwall R;
 XX WPI; 2002-401155/43.
 DR P-PSDB; AAE26349.
 XX
 PT New transgenic non-human mammal that produces detectable levels of a
 PT native human HER2 protein in its mammary gland cells, useful as tumor
 PT models for testing HER2-directed cancer therapies, and for identifying
 PT anticancer agents
 PS Example 2; Page 24-26; 83pp; English.
 XX
 CC The invention relates to a transgenic non-human mammal that produces in
 CC its mammary gland cells detectable levels of a native human HER2 protein
 CC or its fragment. The transgenic animals are useful as tumour models for
 CC testing HER2-directed cancer therapies, and for identifying anticancer
 CC agents. The animals may also be used as source of cells which can be
 CC immortalised in culture, in screening for compounds that have potential
 CC as prophylactic or therapeutic treatments of diseases or disorders
 CC involving expression of HER2. The anti-cancer molecules are useful for
 CC inducing apoptosis or cell death of cancer cells. The present sequence
 CC is human HER-2 cDNA.
 XX
 SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.48e-226 Length: 3768
 Score: 4892.00 Matches: 918
 Percent Similarity: 73.15% Conservative: 0
 Best Local Similarity: 73.15% Mismatches: 1
 Query Match: 96.34% Indels: 336
 DB: 24 Gaps: 1
 SEQ6 (1-919) x AAD43935 (1-3768)

Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGAGATTGCTGGCTCCAAGAAATCTTTGGGAGCCTGGGCATTTCTGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGATGGGACCCAGAGCTCCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAAGTGTTT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGACAGCCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGCTCTTCCAGAACTGCAAGTAATCCGGGGAGCAATCTTCGCAATGGCGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTTCGCTGACCTTGCAGGCGCTGGGCATCAGCTGGCTGGGCTGGCCCTCACATGAGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAGCTCTTTTCGAAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500
Qy	501	GluAspGlyCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGATGTGTGGCGAGGGCTGGCTGCCACCAAGCTGTGCGCCCGAGGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAspCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGCTCCAGGGCCACCAGTGTGTCTCACTCAGACCAAGTTCCTTCGGGGCCAGGAGTGC	1620
Qy	541	ValGluGlyCysArgValLeuGlnGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGAAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACGTG	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCGGTGCCACCTTGAGTGTACGCCCCCAAGTGGCTCAGTGACCTGTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCACTGTGTGGCTGTGCCACTATAGGACCCCTCCCTTCTCGCTGGCCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTAAACCTGACCTCTCCTACATGCCATCTGGAAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCCAGCCTTGCCCCATCACTGCACCCACTCCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCCCGGAGCAGAGAGCCAGCCCTGTGACGTCCATCGCTCTCGCGGTGGTTGGC	1980
Qy	653	-----	653
Db	1981	ATTCTGCTGCTGTGGTCTTTGGGGTGGTCTTTGGGATCTCTCATCAAGCAGCGCAGCAG	2040
Qy	653	-----	653
Db	2041	AAGATCCGGAATACAGATGCGGAGACTGCTGCAGGAAACGGAGCTGTGTGGACCGCTG	2100
Qy	653	-----	653
Db	2101	ACACCTAGCGGAGGATGCCCAACACAGGCGCAGATGCGGATCTCTGAAGAGACGGAGCTG	2160
Qy	653	-----	653

Db	2161	AGGAAGTGAAGTGCTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220
QY	653	-----	653
Db	2221	CCTGATGGGAGAAATGTGAAAATTCAGTGGCCATCAAAGTGTGAGGAAAACACATCC	2280
QY	653	-----	653
Db	2281	CCCAAGCCACAACAAGAAATCTTAGACGAAGCATACGTGATGCTGGTGTGGCTCCCA	2340
QY	653	-----	653
Db	2341	TATGTCCTCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400
QY	653	-----	653
Db	2401	ATGCCCTATGGCTCTCTTAGACCATGTCCGGGAAAACCGCGGAGCTGGGCTCCAG	2460
QY	653	-----	653
Db	2461	GACCTGCTCAACTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGGG	2520
QY	653	-----	653
Db	2521	CTCGTACACAGGACTTGGCCCTCGGAACGTGCTGTTCAAGAGTCCCAACCATGTCAAA	2580
QY	653	-----	653
Db	2581	ATTACAGACTTCGGGTGCTCGGCTGTGGACATTGACGAGACAGATACCATGCAGAT	2640
QY	653	-----	653
Db	2641	GGGGCAAGTGGCCATCAAGTGGATGGCGCTGGAGTCCATTCTCGCGCGGTTCAAC	2700
QY	653	-----	653
Db	2701	CACCAGATGATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGCG	2760
QY	653	-----	653
Db	2761	AAACCTTAGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG	2820
QY	653	-----	653
Db	2821	CTGCCCGAGCCCGCATCTGCACCATTTGATGCTTACATGATCATGTCATAATGTTGGATG	2880
QY	653	-----	653
Db	2881	ATTGACTCTGAATGCGGCCAAGATTCCGGGAGTTGGTCTCTGAATTTCCCGCATGGCC	2940
QY	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	2940
Db	2941	AGGGACCCCGCGCTTTGTGTCTATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTG	3000
QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGTGATGCT	3060
QY	685	GluGlyTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGGAGTATCTGTATCCCGCAGAGGGCTTCTTCTGTCCAGACCTGCCCGGGCGCTGGG	3120
QY	705	GlyMetValHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr	724
Db	3121	GGCATGGTCCACACAGGACCGCAGCTCATCTACAGGAGTGGCGGTGGGACCTGCACA	3180
QY	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	3181	CTAGGCTGGAGCCCTCTCAAGAGGAGGCCCCAGGCTCTCCACTGGCACCTCCGAAGGG	3240
QY	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764

Db	3241	CCTGGCTCCGATGTATTGATGGTGAATGGGGCAGCCAAAGGGGCTGCAAGC	3300
QY	765	LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu	784
Db	3301	CTCCCCACATAGACCCAGCCCTCTACAGCGGTACAGTAGAGACCCACACATACCCCTG	3360
QY	785	ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal	804
Db	3361	CCCTCTGAGACTGATGGCTACGTTGCCCCCTGACCTGCAGCCCCCAGCCTGAATATGTG	3420
QY	805	AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla	824
Db	3421	AAACCCAGCAGATGTTCCGCCCGCAGCCCTTCGCCCGAGAGGGCCCTTGCTGCTGCC	3480
QY	825	ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal	844
Db	3481	CGACCTGCTGGTGCCACTCTGGAAAGGGCCAGACTCTCTCCAGGGGAAGAAATGGGTC	3540
QY	845	VallysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln	864
Db	3541	GTCAAAGACGTTTTTGGCTTTGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCAG	3600
QY	865	GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu	884
Db	3601	GGAGGAGCTGCCCTTCAGCCCACTCTCTCTGCTTCAGCCCGACCTTCACACACCTC	3660
QY	885	TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr	904
Db	3661	TATTACTGGACACGAGCCACAGAGCGGGGGCTCCACCCAGCACCTTCAAGGGGACA	3720
QY	905	ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal	919
Db	3721	CCTACGGCAGAGAACCAGAGTACCTGGGTCTGGAGCTGGCAGTG	3765

Search completed: October 15, 2003, 23:52:57
 Job time : 1038.87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 13234.4 Seconds
(without alignments)
2840.777 Million cell updates/sec

Title: SEQ6
Perfect score: 5078
Sequence: 1 MELAALCRWGLLLALLPPGA.....TFRGTPTAENPEYLGLDVPV 919

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/P/HOLLERAN480/runat_15102003_131912_20536/app_query.fasta_1.4685
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGN_1_1_22724-runat_15102003_131912_20536 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NGG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

SUMMARIES				
Result No.	Score	Query Match %	Length	DB ID
1	4900	96.5	3768	6 AR034479
2	4900	96.5	3768	6 AX201817
3	4900	96.5	3768	6 AX380923
4	4900	96.5	3768	6 AX384604
5	4900	96.5	3768	6 AX465456
6	4900	96.5	4473	6 AR080259
7	4900	96.5	4473	6 AR167390
8	4900	96.5	4473	9 HSERB2R
9	4892	96.3	3768	6 AX060704
10	4892	96.3	3768	6 AX467229
11	4892	96.3	3768	6 AX481438
12	4892	96.3	4530	6 AR202597
13	4892	96.3	4530	6 AR283481
14	4892	96.3	4530	6 AX282577
15	4892	96.3	4530	6 AX587649
16	4892	96.3	4530	6 AX644071
17	4892	96.3	4530	6 BD005474
18	4892	96.3	4530	6 121124
19	4892	96.3	4530	6 159745
20	4892	96.3	4530	9 HUMHER2A
21	4892	96.3	9274	6 AX060703
22	4722	93.0	3678	6 AX505114
23	4438	87.4	3780	4 AB008451
24	4309.5	84.9	2763	6 AX380942
25	4309.5	84.9	2781	6 AX380944
26	4142.5	81.6	4062	10 HAMNEU
27	4138.5	81.5	4727	10 AY116182
28	4134	81.4	3955	6 AX380924
29	4134	81.4	3955	6 121129
30	4134	81.4	3955	6 159750
31	4134	81.4	3955	10 RNNEUR
32	4125.5	81.2	3771	6 AX189662
33	4125.5	81.2	3771	6 AX380925
34	4112	81.0	4694	10 BC046811
35	4112	81.0	4695	10 BC053078
36	3632	71.5	2385	6 AR082744
37	3632	71.5	2385	6 AR099963
38	3632	71.5	2385	6 AR143949
39	3525	69.4	2871	6 AX023363
40	3471	68.4	1872	6 AR181000
41	2585	50.9	1692	6 AX268287
42	2585	50.9	2091	6 AX268288
43	2573.5	50.7	1665	6 AX268285
44	2573.5	50.7	2070	6 AX268286
45	1874.5	36.9	1316	9 AF177761

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1

AR034479	AR034479	Sequence 1 from patent US 5869445.	3768 bp	DNA	linear	PAT 29-SEP-1999
LOCUS	AR034479	Sequence 1 from patent US 5869445.	3768 bp	DNA	linear	PAT 29-SEP-1999
DEFINITION	Sequence 1 from patent US 5869445.	Sequence 1 from patent US 5869445.	3768 bp	DNA	linear	PAT 29-SEP-1999
ACCESSION	AR034479	Sequence 1 from patent US 5869445.	3768 bp	DNA	linear	PAT 29-SEP-1999
VERSION	AR034479.1	GI:5950084	3768 bp	DNA	linear	PAT 29-SEP-1999
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	Unclassified.					
AUTHORS	1 (bases 1 to 3768)					
TITLE	Cheever M.A. and Disis, M.L.					
JOURNAL	Methods for eliciting or enhancing reactivity to HER-2/neu protein					
FEATURES	Patent: US 5869445-A 1 09-FEB-1999;					
source	Location/Qualifiers					
	1..3768					
BASE COUNT	759 a 1171 c 1119 g 719 t					
ORIGIN	/organism="unknown"					
Alignment Scores:						
Pred. No.:	5,39e-187	Length:	3768			
Score:	4900.00	Matches:	919			
Percent Similarity:	73.23%	Conservative:	0			
Best Local Similarity:	73.23%	Mismatches:	0			
Query Match:	96.49%	Indels:	336			
DB:	6	Gaps:	1			
SEQ6 (1-919) x AR034479 (1-3768)						
QY	1 MetCluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuProGlyAla 20					
Db	1 ATGAGCTGGCGGCTTGCCGCTGGGGCTCTCTCGCCCTTTGCCCGCGGAGCC 60					
QY	21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40					
Db	61 GCGAGCACCAAGTGTGCACGGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 120					
QY	41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnClyCysGlnValValGlnGlyAsnLeu 60					
Db	121 ACCCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGPGTGCAGGGAACACTG 180					
QY	61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80					
Db	181 GAATCACTACCTGCCACCACCAATGCCAGCCTGTCTCTTCGCAGGATATCCAGGAGGTG 240					
QY	81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100					
Db	241 CAGGGCTACGTGCTCATCGCTCAACCAAGTCAGGCAGGTGCCACTGCAGAGGCTCGCG 300					
QY	101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120					
Db	301 ATTGTGGAGGACCCAGCTTTTGGAGGCAACTATGCCCTGGCCGCTGTAGACAATGGA 360					
QY	121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140					
Db	361 GACCCGCTGAACAATACACCCCTGTACAGGGGCTCCCGAGGAGGCTGCCGGAGCTG 420					
QY	141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160					
Db	421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCACG 480					
QY	161 LeuCysTyrGlnAspThrIleLeuTrpIlyAspIlePheHisLysAsnAsnGlnLeuAla 180					
Db	481 CTCTGCTACCAGGACACAGATTTTGTGAAGGACATCTTCCACAAGAACAACACCTGGCT 540					
QY	181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200					
Db	541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAA 600					
QY	201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220					
Db	601 GGTCTCCGCTGCTGGGAGAGATTCGAGATGTTCAGAGCCTTCAGCGCCTGACGCGACTGTCTGT 660					

Qy	221	AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluInCys	240
Db	661	GCGGGTGGCTGTGCGCGCTCAAGAGGGCCACTGCCACTGCTGCCATGAGCAGTGT	720
Qy	241	AlaAlaClyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	721	GCTGCCGGCTGTCACGGGGCCCAAGACTCTGACTGCGCTGGCTGCCCTTCAACCA	780
Qy	261	SerGlyTleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATCTGTGAGTGCACTGCCAGCCCTGGTCACTTACAACAGACAGCTTTGAG	840
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTACTGCTGTCC	900
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACCTACCTTTCTACGGAGCTGGGATCTGCACCTCTGCTGCCCCCTGCACAACCA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACACAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGCTGGGCATGGAGCATTCGAGAGGTGAGGCGAGTTACCACTGCCCAAT	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGAGGTTTCTGTGCTCAAGAGATCTTTGGAGCCTGGCATTTCTGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGATGGGAGCCAGCCTCCACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGTTT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGATACAGGTTACTATACATCTCAGCATGGCGGAGCAGAGCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACTCAGCGCTCTTCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACAATGGCGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTCGAAGGGCTGGGCATCAGCTGGCTGGGGTGCCTCACTGAGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGCAGTGGACTGGGCCCTCATCCACCAATAACACCCACCTCTGCTTCGTGCACACGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAGCTCTTTGCGAAACCCGACCAAGCTCTGCTGCCACACTGCCAACCCGCCA	1500
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGTGTGTGGCGAGGGCTGGCTGCCACCACTGTGGCGCCGAGGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGTCCACGGCCACCAGTGTGTCAACTGCAGCCAGTTCTCTTCGGGCCAGGAGTGTC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGAGGAATGCGGAGTACTGCAAGGGCTCCCCAGGGGAGTATGTGTGAATGCAGGCACCTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCGGTGCCACCTGTAGTGTGAGCCCAAGTGGCTAGTGAGCTCTGTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys	600

Db	1741	GCTGACCAAGTGTGGCGCTGTGCCCACTATAAAGACCCCTCCCTTCTGGTGGCCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCCAGCGGTGGAACCTGACCTCTCTACATGCCATCCCTGGAAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Db	1861	GGCCGATGCCAGCCTTGCCCNCACTGACACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProLaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGGGGTGGTGGC	1980
Qy	653	-----	653
Db	1981	ATTCTGCTGTCGTGGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGACAG	2040
Qy	653	-----	653
Db	2041	AAGATCCGGAAGTACACGATCCGAGACTGCTGCAGGAACGGAGCTGGTGGACCCCTG	2100
Qy	653	-----	653
Db	2101	ACACCTAGCGAGCGATGCCCAACAGCGCGAGATGCGGATCCTGAAAGAGACGGAGCTG	2160
Qy	653	-----	653
Db	2161	AGGAAGTGAAGGTGCTTGGATCTGCGCTTTTGGCACAGTCTTACAAGGCGATCTGGATC	2220
Qy	653	-----	653
Db	2221	CCTGATGGGGAAGTGTGAANAATTCACATGCCCATCAAAAGTGTGACGGAAACACATCC	2280
Qy	653	-----	653
Db	2281	CCCAAGCAACAAGAAATCTTAGACGAAGACATACGTGCTGGTGGTGGCTCCCCA	2340
Qy	653	-----	653
Db	2341	TATGTCCTCCGCTTCTGGGCATCTGCTGACATCCACGCTGCAGTGGTGCACACAGCTT	2400
Qy	653	-----	653
Db	2401	ATGCCCTATGCTGCCTCTTAGACCATGTCCGGGAACCCGCGACCCCTGGGCTCCAG	2460
Qy	653	-----	653
Db	2461	GACCTGCTGAACCTGGTGTATCCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2520
Qy	653	-----	653
Db	2521	CTCGTACACAGGGACTGGCGCTCGGAACGTGCTGTCAAGAGTCCCAACCATGTCAA	2580
Qy	653	-----	653
Db	2581	ATTACAGACTTCGGGCTGCTCGGCTGCTGGACATTCAACGACAGAGATACCATGCAGAT	2640
Qy	653	-----	653
Db	2641	GGGGCAAGGTGCCCATCAAGTGGATGGCGTGGAGTCCATTCTCCCGCGCGGTTCCAC	2700
Qy	653	-----	653
Db	2701	CACCAGAGTGTGTGGAGTGTATGGTGTGACTGTGTGGAGCTGTGACTTTTGGGGCC	2760
Qy	653	-----	653
Db	2761	AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAAAGGGGAGCGG	2820
Qy	653	-----	653

Db	2821	CTGCCCCAGCCCCCATCTGCACCATTGATGCTTACATGATCATGTCATGATGTTGGATG	2880
Qy	653	-----	653
Db	2881	ATTGACTCTGAATGTCGGCCAAAGATTCCGGAGTTGGTGTCTGTAATTCTCCCGCATGGCC	2940
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	2941	AGGACCCCCAGCCCTTTTGGTTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTG	3000
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTTCTACCGCTCCTGCTGAGGACGATGACATGGGGGAGCTGGTGGATGCT	3060
Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyValagly	704
Db	3061	GAGGAGTATCTGGTACCCAGCAGGGCTTCTTGTCTCAGACCTTGCCCGGGCGCTGGG	3120
Qy	705	GlyMetValHisArgHisArgSerSerSerThrArgSerGlyGlyGlyAspLeuThr	724
Db	3121	GGCATGCTCCACCACAGCCAGCTCATCTACAGGAGTGGCGGTGGGACCTGCACA	3180
Qy	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	3181	CTAGGGCTGGAGCCCTCTGAAGAGAGAGGCCCCAGGCTCTCCACTGGCACCTCCGAAGG	3240
Qy	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
Db	3241	CTGGGCTCCCATGATTTTGTGTGACCTGGGAATGGGGCGAGCAAGGGGCTGCAAGC	3300
Qy	765	LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu	784
Db	3301	CTCCCCACACATGACCCCGCCCTCTACAGGGGTACAGTGGAGGCCCCACAGTACCCTG	3360
Qy	785	ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal	804
Db	3361	CCCTCTGAGACTGATGCTACGTTGCCCTGACCTGACCTGACGCCCGCCAGCTGAATATGTG	3420
Qy	805	AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla	824
Db	3421	AACAGCCCATGATGTGGCCCCCAGCCCTTCGCCCGGAGAGGGCCCTCTGCCCTGTGCC	3480
Qy	825	ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal	844
Db	3481	CGACCTGCTGCTGCCACTCTGGAAGAGGCCCAAGACTCTCTCCCAGGGAAGATGGGTC	3540
Qy	845	ValLysaspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln	864
Db	3541	GTCAAGACGCTTTTGGCTTTGGGGTGGCGTGAGAACCCCGAGTACTTGACACCCAG	3600
Qy	865	GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu	884
Db	3601	GGAGGAGTCCCTTACGCCCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3660
Qy	885	TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr	904
Db	3661	TATTACTGGGACCCAGGACCCACAGAGCGGGGCTCCACCAGCACCTTCAAGAGGACA	3720
Qy	905	ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal	919
Db	3721	CCTACGSCAGAGAACCAGAGTACCTGGGTCTGGACGTGCCAGTG	3765

RESULT 2

AX201817	LOCUS	AX201817	3768 bp	DNA	linear	PAT 30-AUG-2001
AX201817	DEFINITION	Sequence 1 from Patent WO0153463.				
AX201817	ACCESSION	AX201817				
AX201817.1	VERSION	AX201817.1	GI:15391666			
	KEYWORDS					
	SOURCE	Homo sapiens (human)				
	ORGANISM	Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				

1 Cheever, M.A. and Hand-Zimmerman, S.
 Compounds and methods for prevention and treatment of her-2/ neu
 associated malignancies
 Patent: WO 0153463-A.1 26-JUL-2001;
 CORIXA CORPORATION (US)
 FEATURES Location/Qualifiers
 source 1..3768
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 CDS 1..3768
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC60126.1"
 /db_xref="GI:15391667"
 /translation="MELAAALCPWGLLLALLPDGAASQTQVCTGTDMLRLPASPETHLD
 MLKHLTGGQVOGNGLELYLPTNASLSFLDQLQIEQGYVLIAHNOVRQVPQLRLRIV
 RQPLFEDNYALAVLDNGDPLNNITPVTGASPGELRELQLRSLTEILKGVLFQRNPRL
 LCYQDTILNKRNDIPKHNQLALITDTRNSRACHPCSPCKMGSGSCWSESSEDQSLVTRT
 VCAAGCARCKGPLTDCHEQCAAGCTGPKHSDCLACLFHNHSGICELHCPALVTVNT
 DTFESNPEGRVTFGASCVATCPYNLSTDVGSCTLVCPFLHNQEWYTABDGTQRCCK
 SPFCARVCVGLGNEHLREVRVNTSANIQEFAGCKKIFGSLAFLPESFDGDPASNTAPL
 QPQLQVFTLEBITGLYLISAWPDSLPDLVSFQNLQVLRILHNGAYSLTLQVLGI
 SWLGRSLRELSEGLALIHHNTHLCFVHTVPMDFNPHQALLHHTANRPEDCEVGGI
 LACHOLCARHGHWGPGTQCVNGSQFLRGQECVEECRVLQGLPREYVNAHRHCLPCHQE
 COPNGSTFCGPEADQVCACAHKDPFPCVARCPSSVKPDLISYMPWKFPDDEGACQ
 PCPINTGHSVDLDDKGCAPORASPLTSLISAVGILLVVLGVVGGILIKRPOOKI
 RKYTMRRLLQTELEPVLPSGAMPNOAQRILKTELKRVKVGSGARGTVYKGLWI
 PDGENKIPVAIKVLRNTSPKANKEITDEAVVMAGVSGPYPSRLIGLCITSTVLQGI
 QLMPYCGLLDHVRNGLRSGDQLNNCKQIAKMSYLEDVRLVHRDLAARNVLVKSP
 NHHVITDGLRLADIDETEHADGGKVTIKWMALESILRRRTHQSDVMSYGVTRLEL
 LMTFGAKYVDIPARITDPLELGERLPOPICTIDVYIMLVKMWIDSECRPFRREL
 VSFESRMARDPQRFVVIQNEIDLGASPLDSTFVSRLLEDMDGLVDASELYLVQOQF
 FCDPAPGAGGVVHHRSSSTRSGGDLTLGLEPSEERAPRSPAPSGAGSDVDFG
 DLGMAAGKLQSLTDPQLQRYSEDPTVPLPSETDGVYAPLTCSPPQPEYVNPQDVR
 POPPSREGPLPKAARPAAGATLERPKTLSPKNGVVKDVFAGFAGVAPENPEYLTPOGGA
 PQPHPPATSPAFDNLVYWDQDPPERGAAPPSTFKGTPTAENPEYLTGLDVPV"
 BASE COUNT 759 a 1171 c 1119 g 719 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,39e-187 Length: 3768
 Score: 4900.00 Matches: 919
 Percent Similarity: 73.23% Conservative: 0
 Best Local Similarity: 73.23% Mismatches: 0
 Query Match: 96.49% Indels: 336
 DB: 6 Gaps: 1
 SEQ6 (1-919) x AX201817 (1-3768)
 Qy 1 MetGluLeuAlaAlaLeuCysArgTgTgLyLeuLeuLeuAlaLeuLeuProGlyAla 20
 Db 1 ATGGAGTGGCGGCTGTGTGCCCTGGGGCTCCTCCTCGCCCTCTGTGCCCCCGAGCC 60
 Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 61 GCAGACACCAAGTGTGCACCGGCACAGACATGAGCTGCGGCTCCTCGCATGTCGCGAG 120
 Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 121 ACCCACCTGACATGCTCCGCCACCTCTACACGGGCTGCCAGTGTGTGCAGGGAACCTG 180
 Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 181 GAATCACCCTACCTGCGCCACCAATGCGAGCTGTCTCTCCTGCGAGGATATCCAGAGG 240
 Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 241 CAGGGCTAGTGTCTCATCGTTCACACCAAGTGGAGCAGGTGCCACTGCAGAGGCTGCGG 300
 Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120

301	ATTGTGGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCCGTGCTAGACAATGGA	360
121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
361	GACCCGCTGAACAATACCAACCCCTGTACAGGGGSCCTCCCCAGGAGSCCTGCGGAGCTG	420
141	GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGln	160
421	CAGCTTCGAAGCCTCACAGATCTTCGAAAGGAGGGTCTTGATCCAGCGAACCCTCCAG	480
161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
481	CTCTGTACCAGACACAGATTTTGTGAAGGACATCTTCACAGAACAACACAGCTGGCT	540
181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
541	CTCACACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTCTCCGATGTGTAA	600
201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
601	GGCTCCCGCTGCTGGGAGAGATTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGT	660
221	AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
661	GCCGGTGGCTGTGCCCGCTCAAGGGGCCACTGCCCACTGACTGCTGCTCATGACAGCTGT	720
241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
721	GCTGCCGGCTGACAGGGCCCCAAGCACTCTGACTGGCTGGCTGCCCTCCACTTCACCCAC	780
261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
781	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACAAACACAGACACGCTTTGAG	840
281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
841	TCCATGCCCAATCCCGAGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTGCC	900
301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
901	TACAACCTACCTTTCTACGGACGTGGGATCTTGACCCCTGCTGTGCCCCCTGACACACCA	960
321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
961	GAGGTGACACACAGAGGATGAACACACACGCGTGTGAGAAGTGCAACAAGCCTGTGCCCGA	1020
341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
1021	GTGTGCTATGCTCTGGGCATGGAGCACTTCGAGAGGTGAGGGCAGTTACCACTGGCCAAT	1080
361	IleGlnIlePheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
1081	ATCCAGAGATTGCTGGCTCGAAGAATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC	1140
381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
1141	TTTGTATGGGACCCAGCCTCAACACTGCCCGCTCCAGCCAGACAGCTCCAAGTGT	1200
401	GluThrLeuGluGluLeuThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro	420
1201	GAGACTCTGGAAGAGATCACAGGTATACCTATACATCTCAGCATGGCGGACAGCCTGCCT	1260
421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
1261	GACTCTAGCGTCTTCCAGAACCTCGCAAGTAAATCCGGGGAGCAATCTGCAACAATGGCGCC	1320
441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTripleuGlyLeuArgSerLeuArgGlu	460
1321	TACTCGCTGACCTCGAAGGCTGGGATCAGCTGGCTGGGGCTGCCCTCACTGAGGAA	1380
461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
1381	CTGGCAGTGGACTGGCCCTCATCCACATTAACACCCACTCTGCTTGTGCACACGGTG	1440

QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
DB	1441	CCCTGGGACCAGCTCTTTCCGAACCGCGACCAAGGCTCTGCTCCACACTGCCAACCGGCCA	1500
QY	501	GlUAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuLeuCysAlaArgGlyHisCys	520
DB	1501	GAGGACGAGTGTGTGGCGGAGGGCTTGGCCCTGCCACACAGCTGTGCGCCGAGGCACTGC	1560
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
DB	1561	TGGGTCCAGGGCCACCAGTGTGTCAACTGCAGCCAGTTCTCTTCGGGGCCAGGAGTGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
DB	1621	GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
DB	1681	TTGCCGTGCCACCTCAGTGTGAGCCCAAGATGGCTCAGTGACCTGTTTGGACCGGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrIysAspProProPheCysValAlaArgCys	600
DB	1741	GCTGACCAAGTGTGGCCCTGTGCCACATATAAGGACCTCCCTTCTGGTGGCCCGCTGC	1800
QY	601	ProSerGlyValIysProAspLeuSerTyrMetProIleTyrPheProAspGluGlu	620
DB	1801	CCGAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspIys	640
DB	1861	GGCGATGCGAGCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
DB	1921	GGTGTCCCGCGGAGCAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGC	1980
QY	653	-----	653
DB	1981	ATTCTGCTGTCGTGTCTTGGGGTGTGTTTGGGATCCTCATCAAGCGACGGCAGCTG	2040
QY	653	-----	653
DB	2041	AAGATCGGGAAGTACACGATCGGAGACTGTGCGAGGAACGGAGCTGGTGGACCCGCTG	2100
QY	653	-----	653
DB	2101	ACACCTAGCGGAGCGATGCCAACACGAGCGCAGATGCGGATCCTGAAGAGACGCGAGCTG	2160
QY	653	-----	653
DB	2161	AGGAAGGTGAAGTGCTTGGATCTGGCGCTTTTGGCACAGCTCAACAAGGGCATCTGGATC	2220
QY	653	-----	653
DB	2221	CCTGATGGGAGAGTGTGAATAATTCAGTGGCCATCAAAAGTTGTAGGGAAAAACATATCC	2280
QY	653	-----	653
DB	2281	CCCAAGCCCAACAAGAAATCTTAGACNAGCATACGTATGGTGTGGTGTGGGCTCCCCA	2340
QY	653	-----	653
DB	2341	TATGTCTCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400
QY	653	-----	653
DB	2401	ATGCCCTATGGTGCCTCTTAGACCATGTCCGGGAAAAACCGGAGCGCTGGGCTCCGAG	2460
QY	653	-----	653
DB	2461	GACCTGCTGACTGGTGTATGCAGATGTGCCAAGGGGATGAGCTTACCTGGAGGATGTCCGG	2520


```
Qy 725 LeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGly 744
|||||
Db 3181 CTAGGCTGGAGCCCTCTGAGAGAGAGGCCCCAGGTCTCCACTGCCACCTCCGAGGG 3240

Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
|||||
Db 3241 GCTGGCTCCGATGATTGATGCTGACCTGGGAATGGGGACGCCAAGGGCTGCAAGC 3300

Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
|||||
Db 3301 CTCCACACATACCCACCCCTCTACAGCGGTACAGTGAGAGACCCACAGTACCCCTG 3360

Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
|||||
Db 3361 CCTCTGAGACTGATGGCTAGCTTGCCCTCCCTGACCTGACGCCCTGATATG 3420

Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
|||||
Db 3421 AACGAGCCAGATGTCGGCCCGCCAGCCCTTCGCCCGAGAGGGCTCTGCTGCTGCC 3480

Qy 825 ArgProAlaGlyAlaThrLeuGluArgProGlyThrLeuSerProGlyLysAsnGlyVal 844
|||||
Db 3481 CGACCTGCTGGTCCCTCTGGAAGAGGCCCAAGACTCTCTCCCGAGGGAAGATGGGTC 3540

Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
|||||
Db 3541 GTCAAGACGCTTTTGGCTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3600

Qy 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
|||||
Db 3601 GGAGGAGCTGCCCTCAGCCCGCCACCTCTCTCTGCTTCCAGCCAGCTTGCACACCTC 3660

Qy 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
|||||
Db 3661 TATTACTGGACGAGACCCACGAGCGGGGGCTCCACCCAGCACCTTCAAGGGGACA 3720

Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
|||||
Db 3721 CCTACGCGAGAGAACCAGAGTACCTGGGTCTGGAGTGCAGCTG 3765

RESULT 4
LOCUS AX384604
DEFINITION Sequence 1 from Patent WO0214503.
ACCESSION AX384604
VERSION AX384604.1 GI:19577806
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hand-zimmermann,S., Cheever,M.A., Foy,T.M., Lodes,M.J., Kalos,M.D.,
McNeill,P.D. and Vedvick,T.S.
TITLE Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 1 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/feature="unnamed protein product"
/codon_start=1
/db_xref="GI:19577807"
/translation="MELALCRWGLLLALLPPGAASVCTGTDMKRLRPLASPETHLD
MRHLVQGVQVQGNLELYLPTNASLSFLQIQVOGYVLIHAHQVQVPLQRIRIV
RGTLQFEDNIALVDNGDPLNNITPTVGASPGGURELQLRLSTELKKGGLIQNRNQ
LCYQDTILMKDIFHKNNQLALTLDITNRSRACHPCSPMCKGSRGWGESSEDCQSLTRT
```


Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	301	ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATCCCTGGCGGTCTAGACAATGGA	360
	121	AspProLeuAsnAsnThrThrProValThrClyAlaSerProGlyGlyLeuArgGluLeu	140
	361	GACCCGCTGAACAATACACCCCTGTCCAGGGGCTCCCGAGAGGCTTCGGGAGCTG	420
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Db	421	CAGCTTGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCGAG	480
	161	LeuCysTyrGlnAspThrIleLeuTrpPlyAspIlePheHisLysAsnAsnGlnLeuAla	180
	481	CTCTGCCTACCAAGGACACGATTTTGTGAAGGACATCTCCACAAGAACAACACAGCTGGCT	540
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	541	CTCACACTGATAGACACCAACCGCTCTCGGGCTTCACCACCTGTTCTCCGATGTGTAAG	600
	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
	601	GGTCCCGCTGTGGGAGAGATTCGAGGATGTCTAGAGCTTGACGCCGACTGCTGT	660
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	661	GCCGCTGGCTGTGCCCTCTCAAGGGGCCACTGCCACTGACTGCTGCCATGAGCAGTGT	720
	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
	721	GCTGCCGGCTGCAGGGCCCCAAGCACTCTGACTGCCCTGGCTGCCCTTCAACCCAC	780
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATCTGTGAGCTGGACTGCCAGCCCTGGTCACTACACACAGACAGCTTTGAG	840
	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
	841	TCCATGCCCAATCCGAGGGCCGATATACATATCGGCGCCAGCTGTGTGACTGCTGTCCC	900
Qy	301	TyrAsnThrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACCTACCTTCTACGGACGTGGGATCTGCACCCCTGCTGTGCCCTTCACACACAA	960
	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
	961	GAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGCTCTGGGCATGGAGCACTTCGAGAGGTGAGGGCAGTTACCACTGCCAAT	1080
	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
	1081	ATCCAGAGATTTGCTGGCTCCAGAAGATCTTTGGAGCCTTGGCATTTCTGCCCGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400
Db	1141	TTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGACAGCTCCCAAGTGT	1200
	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
	1201	GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCGGAGACGCTGCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTACGCTCTTCCAAGACCTTCAAGTAATCCGGGGAGCAATCTGCACAATGGCGCC	1320
	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
	1321	TACTCGCTGACCTTGAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCGCTCCTCCTGAGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480

1381	Db		CTGGCAGTGGACTGGCCCTCATCCACATTAACACCCACCTCTGCTTCTGTCGACACGGTG	1440
481	Qy		ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
1441	Db		CCCTGGGACAGACTCTTTCCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500
501	Qy		GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
1501	Db		GAGGACAGTGTGTGGCGAGGGGCTGGCTGCCACCACTGTGCGCCCGAGGCACTGC	1560
521	Qy		TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
1561	Db		TGGGGTCCAGGGGCCACCCAGTGTGTCACTGACGACAGTCTCTTCTGGGGCCAGGAGTGC	1620
541	Qy		ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
1621	Db		GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGTAATGCCAGCACTGT	1680
561	Qy		LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
1681	Db		TTGCCGTGCCACCTCTAGTGTCAAGCCCAAGATGGCTCAGTGAACCTGTTTGGACCGGAG	1740
581	Qy		AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
1741	Db		GCATACCAAGTGTGGGCTGTGCCCATATAAGACCCCTCCCTTCTGCGTGGGCCCGCTGC	1800
601	Qy		ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
1801	Db		CCCAGCGGTGGAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
621	Qy		GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
1861	Db		GGCGCATGCCAGGCTTGGCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
641	Qy		GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
1921	Db		GGTGTCCCCCGCAGAGAGAGCCGCGCTCTGACGTCCATCATCTCTGCGGTGGTTGGC	1980
653	Qy		-----	653
1981	Db		ATTCTGTGTCGTGTGCTTTGGGGTGGTCTTTTGGGATCTCTCATCAAGCAGCGCAGCAG	2040
653	Qy		-----	653
2041	Db		AAGATCCGGAAGTACAGGATGCGGAGACTGCTGCAGGAACAGGAGCTGTGTGGAGCCGCTG	2100
653	Qy		-----	653
2101	Db		ACACCTAGCGGAGCGATGCCCAACCAAGCCGACAGTGGGATCTCTGAAAGAGACGGAGCTG	2160
653	Qy		-----	653
2161	Db		AGGAAGGTGAAGTGTGTGATCTGGCGCTTTTGGCACAGTCTACAAGGCACTCTGGATC	2220
653	Qy		-----	653
2221	Db		CCTGTATGGGAGAAATGTGAAAATCCAGTGGCCATCAAAAGTTTGAGGGAAAAACATCC	2280
653	Qy		-----	653
2281	Db		CCCAAGCCCAAGAAATCTTAGACGAGCATACGTGTATGGCTGTGTGGGCTCCCCA	2340
653	Qy		-----	653
2341	Db		TATGTCTCCGCGCTTCTGGGCATCTGCCGTGACATCCACGGTGCAGCTGTGTGACACAGCTT	2400
653	Qy		-----	653
2401	Db		ATGCCCTATGGCTGCCCTCTTAGACCATGTCCGGGAAAAACCGCGGACGCCCTGGCTCCCAAG	2460
653	Qy		-----	653

```
Db 2461 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2520
QY 653 -----
Db 2521 CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA 2580
QY 653 -----
Db 2581 ATTACAGACTTCGGGCTGGCTCGGCTGTGGACATTGACGAGACAGTACCATTGCAGAT 2640
QY 653 -----
Db 2641 GGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCGCGCGGGGTTCACC 2700
QY 653 -----
Db 2701 CACCAGAGTATGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2760
QY 653 -----
Db 2761 AAACCTTACGATGGATCCAGCCGGGAGATCCCTGACCTGTGTGAAAGGGGAGCGG 2820
QY 653 -----
Db 2821 CTGCCCCAGCCCCCATCTGCACATTGATGTCTACATGATCATGGTCAAAATGTTGGATG 2880
QY 653 -----
Db 2881 ATTGACTCTGAATGTCCGCCAAGATTCGGGGAGTGGTGTCTGTAATTCCTCCGCATGGCC 2940
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 2941 AGGGACCCCGAGCGCTTGTGTCTATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTG 3000
QY 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGACCTTCTACCGCTCACTGTGTGGAGACGATGACATGGGGACCTGTGTGATGT 3060
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGGTACCCAGCAGGCTTCTCTGTCCAGACCTGCCCCGGCGGTGGG 3120
QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
Db 3121 GGCATGGTCCACACAGCCAGCCAGCTCATCTACCAGGAGTGGCGGTGGGACCTGCACA 3180
QY 725 LeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3181 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCAGGTCTCCACTGGCACCCCTCCGAAGG 3240
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3241 GCTGGCTCGATGATTTCATGGTGACCTGGGAATGGGGGCGCAGCCAAAGGGGTGCAAGC 3300
QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerSerGluAspProThrValProLeu 784
Db 3301 CTCCCCACACATGACCCCGCTCTACAGCGGTACAGTGAGGACCCACAGTACCCTCG 3360
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3361 CCTCTGAGACTGATGGCTACGTGGCCCCCTTGACCTGACCGCCCGAGCTGAATATGTC 3420
QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3421 AACCACCCAGATGTTCCGGCCCCAGCCCCCTTCGCCCGAGAGGCCCTCTGCTGCTGCC 3480
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3481 CGACCTGCTGGTGCCACTCTGGAAGGCCCAAGACTCTCTCCCGAGGGAAGATGGGGTC 3540
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3541 GTCAAGAGCTTTTGGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3600
```

```
QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3601 GGAGGAGCTGCCCTCAGCCCCACCTCTCTCTGCTTCCAGCCAGCTTCCGACAACTC 3660
QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3661 TATTACTGGGACGAGGACCCACGAGCGGGGGGTCCACCCAGCACCTTCAAAGGGACA 3720
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3721 CCTACGGCAGAACCCAGAGTACCTGGGTCTGGAGCTGCCAGTG 3765

RESULT 6
LOCUS AR080259 4473 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 1 from patent US 5968748.
ACCESSION AR080259
VERSION AR080259.1 GI:10006994
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4473)
AUTHORS Bennett,C.Frank., Lipton,A. and Witters,L.M.
TITLE Antisense oligonucleotide modulation of human HER-2 expression
JOURNAL Patent: US 5968748-A 1 19-OCT-1999;
FEATURES
source I. .4473
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:
Pred. No.: 6,37e-187 Length: 4473
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 6 Gaps: 1

SE06 (1-919) x AR080259 (1-4473)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGCTGGCGCCTGTGCGCTGGGGGCTCTCTCTGCTCTTGTCCCGGAGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProLysSerProGlu 40
Db 235 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGTCGGCTCTCCCTGCCAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCCACTGGAGCATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGGTGGAGGAAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 355 GAACCTCACCTACTGCCCAACCAATGCCACCTGTCTCTCTCCAGGATATCCAGGAGGTG 414
QY 81 GlnGlyTyrValLeuIleAlaHisGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGGCTAGTGTCTCATCTCACACCAAGTGAGGACAGGTCCCACTGCAGAGGCTGCGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCTGGCTAGCAATGGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGAACAATACCACTCCCTGTCACAGGGGCTCCCAAGGAGGCTGCGGGAGCTG 594
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
```

Db	595	CAGCTTCGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAACCCCCAG	654	Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Qy	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180	Db	1735	TGGGGTCAGGGCCACCCACAGTGTCAACTGCAGCCAGTTTCTTCGGGGCCAGGAGTGC	1794
Db	655	CTCTGCTACCAAGGACACGATTGTGTGAAGAGACATCTTCCCAAGAACAACCACTGGCT	714	Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200	Db	1795	GTGGAGGATGCCGAGTACTGCAGGGGTCTCCAGGGAGTATGTGAATGCCAGGCACACTGT	1854
Db	715	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGACTGTGAAG	774	Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Qy	201	GlySerArgCysTrpGlyLysSerGluAspCysGlnSerLeuThrArgThrValCys	220	Db	1855	TTGCGGTGCCACCCCTGAGTGTACGCCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG	1914
Db	775	GGCTCCCGCTCTCGGGAGAGATCTGAGGATTGTACAGCCCTCAGCGCACGTGCTGT	834	Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGluCys	240	Db	1915	GCTGACCAGTGTGGGCTGTGCCCACTATTAAGACCCCTCCCTTCTCGGTGGCCGCTGC	1974
Db	835	GCCGGTGCTGTGCCGTGCAAGGGCCACTGCCACTGACTGCTGCCATGAGCAGTGT	894	Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	Db	1975	CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTCGAAAGTTTCCAGATGAGGAG	2034
Db	895	GCTGCGGCTGCACGGGCCCAAGCACCTCTGACTGCTGCGCTGCCTCCACTTCAACAC	954	Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	Db	2035	GGCCCATGCCAGCCTTGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	2094
Db	955	AGTGGCATCTGTGAGCTGCACCTGCCACGCCCTGTGTCACTTACAACACAGACACGTTTGAG	1014	Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	Db	2095	GGCTGCCCGCCGAGCAGAGACCGCCCTCTGACGTCCATCATCTCTGCGGTGGTGGC	2154
Db	1015	TCCATGCCCATCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTGCC	1074	Qy	653	-----	653
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	Db	2155	ATTCTGTGTGTCGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGCAGCGCAG	2214
Db	1075	TACAACTACCTTTCTACGAGCGTGGGATCTCTGCACCCCTCGTCTGCCCTTCCACACCAA	1134	Qy	653	-----	653
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	Db	2215	AAGATCCGGAAGTACACGATCGGAGACTGCTGCGAGAAACGGAGCTGGTGGAGCCGCTG	2274
Db	1135	GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCACAAGCCCTGTGCGCGA	1194	Qy	653	-----	653
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	Db	2275	ACACCTAGCGGAGCGATGCCCAACCGCGCAGATCGGGATCTCTGAAGAAGACGGAGCTG	2334
Db	1195	GTGTGCTATGCTTGGGCATGGAGCATTGCGAGAGGTGAGGCGCATTTACCAAGTGGCCAAT	1254	Qy	653	-----	653
Qy	361	IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	Db	2335	AGGAAGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2394
Db	1255	ATCCAGGAGTTTGTGGTGCACAAGATCTTTGGAGCCCTGGCATTTCTGCCGGAGAGC	1314	Qy	653	-----	653
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400	Db	2395	CCTGATGGGAGAAATGTGAAAATTTCCAGTGGCCATCAAAGTGTGTGAGGGAAAAACATCC	2454
Db	1315	TTTGATGGGAGCCAGCCTCCAAACACTGCCCGCTCCAGCCAGCAGCAGCTCCAAGTGT	1374	Qy	653	-----	653
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro	420	Db	2455	CCCAACCCCAACAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGGGCTCCCCA	2514
Db	1375	GAGACTCTGGAAAGATCACAGTTACTATACATCTCAGCATGCGCGGAGCAGCCTGCCT	1434	Qy	653	-----	653
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	Db	2515	TATGCTCTCCCGCTTCTTGGGCATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2574
Db	1435	GACCTCACGCTCTCCAGAACCTCAAGTAATCCGGGACCAATCTGCACAATGGCGCC	1494	Qy	653	-----	653
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460	Db	2575	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGAAAAACCGGACGCGCTGGGCTCCAG	2634
Db	1495	TACTCGCTGACCTTCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGAA	1554	Qy	653	-----	653
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480	Db	2635	GACCTGCTGAACCTGGTGTATGACAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2694
Db	1555	CTGGGAGTGACTGGCCCTCATCCACATACACCACCTCTCTGCTCGTGCACACCGTG	1614	Qy	653	-----	653
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	Db	2695	CTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAAA	2754
Db	1615	CCCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGTCTCCACACTGCCAACCGGCCA	1674	Qy	653	-----	653
Qy	501	GluAspGluCysValGlyLeuGlyLeuAlaCysHisGlnLeuCysAlaArgHisCys	520	Db	2755	ATTACAGACTTTCGGGCTGGCTCGGCTGCTGGACATTTCACGAGACAGAGTACCATGCAGAT	2814
Db	1675	GAGCAGAGTGTGTGGGAGGGCTGCGCTGCCACCAAGCTGTGCGGCCGAGGGCAGCTGC	1734				


```
QY 653 ----- 653
Db 2815 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGTTTCACC 2874
QY 653 ----- 653
Db 2875 CACCAGAGTGTGTGGAGTTATGTGTCACTGTGTGGAGCTGATGACTTTTGGGGCC 2934
QY 653 ----- 653
Db 2935 AAACCTTAGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2994
QY 653 ----- 653
Db 2995 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTCAAATGTTGGATG 3054
QY 653 ----- 653
Db 3055 ATTGACTCTGAATGTCGGCCAAAGATTCCGGGAGTTGGTGTCTGAATTCCTCCGCAATGGCC 3114
QY 654 ----- 654
Db 3115 AGGGACCCCGCGCTTTGGTCTCATCCAGATGAGACTTGGGCCCCAGCCAGTCCCTTG 3174
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3175 GACAGCACCTTCTACCGCTCACTGTGTGGAGGACGATGACATGGGGACCTGTGGATGCT 3234
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3235 GAGGAGTATCTGGTACCCAGCAGGCGCTTCTTGTCTCAGAGCCCTGCCCGCGCGCTGGG 3294
QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
Db 3295 GGCATGGTCCACCACAGGACCGCAGCTCATCTACAGGAGTGGCGGTGGGACCTGAC 3354
QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3355 CTAGGCTGGAGCCCTCTCAAGAGGAGGCCCGCCAGCTCTCCACTGSCACCTCCGAAGG 3414
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeuGlnSer 764
Db 3415 GCTGGCTCGATGTATTTGATGGTACCTGGGAATGGGGCGACGCCAAGGGGCTGCAAGC 3474
QY 765 LeuProThrHisAspProSerProLeuGlnAlaTyrTyrSerGluAspProThrValProLeu 784
Db 3475 CTCCCCACATGATACCCAGCCCTCTACAGCGGTACAGTACAGACCCACAGTACCCCTG 3534
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3535 CCTCTGAGACTGATGGCTACGTTGGCCCCCTGACCTGCAGCCCGCCAGCCTGAATATGTG 3594
QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3595 AACACGCCAGATGTTGGCCCGCCAGCCCTCTGCCCCGAGAGGGCCCTGCTGCTGTGCTCC 3654
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3655 CGACCTGCTGGTGGCCACTTGAAAGGCCACAGACTCTCCCCAGGGAAGATGGGGTC 3714
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3715 GTCAAGAGAGTTTTCCTTTGGGGTGGCGTGGAGAACCCCGAGTACACACCCAG 3774
QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3775 GGAGGAGCTGCCCTCAGCCCGCCAGCCCTCTCTGCTGCTTCCAGCCCGCTTCCACAACTTC 3834
QY 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3835 TATTACTGGACGAGACCCAGCAGCGGGGGCTCCACCCAGACCTTCAAGGGGACA 3894
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
```

```
Db 3895 CCTAGGCGAGAACCCAGAGTACCTGGCTGGAGTCCAGTG 3939
RESULT 7
LOCUS AR167390 4473 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 26 from patent US 6287569.
ACCESSION AR167390
VERSION AR167390.1 GI:17903168
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4473)
AUTHORS Kips, T.J. and Wu, Y.
TITLE Vaccines with enhanced intracellular processing
JOURNAL Patent: US 6287569-A 26 11-SEP-2001;
FEATURES Location/Qualifiers
source 1..4473
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:
Pred. No.: 6,37e-187 Length: 4473
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 6 Gaps: 1
SEQ6 (1-919) x AR167390 (1-4473)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCTCTCGCCCTCTTGTCCCGGAGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCAAGTGTGCACCGCACACATGAAGCTCGGCTCCCTGCCAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCACCTGGACATGCTCGCCACCTCTACAGGGCTGCAGGTGTGTCAGGAAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 355 GAACCTACCTACTGCCCAATGCCAGCTGTCTTCTTCAGGATATCCAGGAGGTG 414
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGGCTAGTGTCTCATCTCACACCAAGTGCAGCAGTCCACTGCAGAGGCTGCGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCTGCTGCTGCTAGACATGGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGAACAATACCACTCCAGGGGCTTCCAGGAGGCTCCCGAGGAGCTGCGGAGCTG 594
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTCGAAGCTCACAGAGATCTTGAAGGAGGGGTCTTGTATCCAGCGGAACCCCGAG 654
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db 655 CTCTGCTACCGAGACACGATTTTGTGGAAGGACATCTTCCACAAGAAACAACAGCTGGCT 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACACTGATAGACACCAACCGCTCTCGGGGCTGCCACCCCTGTCTCCGATGTGTAAAG 774
```



```

Db      2935 AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2994
QY      653 -----
Db      2995 CTGCCCCAGCCCCCATCTGCACCATTTGATGCTACATGATCATGTCGTAATGTTGGATG 3054
QY      653 -----
Db      3055 ATTGACTCTGAATGTCGCCAAGATTCGGGAGTTGGTCTGTAATTCCTCCCGCATGGCC 3114
QY      654 -----
Db      3115 AGGACCCCGCAGCGCTTTGTGTCATCCAGATGAGGACTTGGCCGACCCAGTCGCTTGG 3174
QY      665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db      3175 GACAGCACTTCTACCGCTCACTGCTGGAGGAGCATGACATGGGGACCTGGTGGATGCT 3234
QY      685 GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly 704
Db      3235 GAGGAGTATCTGTATCCAGCAGGCGCTTCTTCTGTCACAGACCTGCCCGCGCGCTGGG 3294
QY      705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
Db      3295 GGCATGGTCCACACAGGACCGACGCTCATCTACAGGAGTGGCGGTGGGACCTGAC 3354
QY      725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db      3355 CTAGGCTGGAGCCCTCTGAGAGGAGGCCCGCAGGTCTCCACTGCGACCTCCGAGGG 3414
QY      745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db      3415 GCTGGCTCGATGATTTGATGTCACCTGGGAATGGGGCAGCCCAAGGGCTGCAAGC 3474
QY      765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db      3475 CTCGCCACATGACCCCGCCCGCTCTACAGCGGTACAGTGAGACCCCGACATACCCCG 3534
QY      785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db      3535 CCCTCTGAGACTGATGGCTACGTTGCTCCCGCTGACCTGCAGCCCCCAGCCCTGAATATG 3594
QY      805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db      3595 AACACCCAGATGTTCCGCCCGCAGCCCGCTTCCGCCGAGAGGCGCTCTGCTGCTGCTG 3654
QY      825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db      3655 CGACCTGCTGGTGCCACTCTGGAAGGCCACAGACTCTCTCCCGAGGGAAGTGGGGTC 3714
QY      845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db      3715 GTCAAGACGTTTTTGCCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3774
QY      865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db      3775 GGAGGAGCTGCCCTCAGCCCGCACCCTCTCTCTGCTTCCAGCCCGCAGCCCTTCGACACCTC 3834
QY      885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db      3835 TATTACTGGACGACGACCCAGACCGAGCGGGGGCTCCACCCAGCACCTTCAAGGGGACA 3894
QY      905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db      3895 CCTACGCGACAGAACCCAGAGTACCTGGTCTGGAGTCCAGTG 3939
Db      3939 -----
RESULT 8
HEADER  HSERB2R
LOCUS   HSERB2R
DEFINITION Human c-erb-B-2 mRNA.
ACCESSION X03363
VERSION X03363.1 GI:31197
KEYWORDS cell surface glycoprotein; cellular oncogene; erb-B-2 cellular;

```

```

glycoprotein; growth factor receptor; kinase; neu cellular
oncogene; transmembrane protein; tyrosine kinase.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4473)
AUTHORS Yamamoto,T., Ikawa,S., Akiyama,T., Samba,K., Nomura,N.,
Miyajima,N., Saito,T. and Toyoshima,K.
TITLE Similarity of protein encoded by the human c-erb-B-2 gene to
epidermal growth factor receptor
JOURNAL Nature 319 (6050), 230-234 (1986)
MEDLINE 86118663
PUBMED 3003577
REFERENCE 2
AUTHORS Papewalis,J., Nikitin,A.Yu. and Rajewsky,M.F.
TITLE G to A polymorphism at amino acid codon 655 of the human
erbB-2/HER2 gene
JOURNAL Nucleic Acids Res. 19 (19), 5452 (1991)
MEDLINE 92020265
PUBMED 1681519
COMMENT The c-erb-B-2 protein shows similarity to the epidermal growth
factor receptor.
FEATURES
Location/Qualifiers
1..4473
/organism="Homo sapiens"
/mol_type="mRNA"
/strain="cell line MKN-7"
/db_xref="taxon:9606"
175..3942
/codon="c-erb-B-2 precursor"
/codon_start=1
/protein_id="CAA27060.1"
/db_xref="GI:31198"
/db_xref="SWISS-PROT:P04626"
/translation="MELALCRWGLLLALLPPGAASQTQCTGDMKRLPASPETHLD
MURHYQGVVQVGNLELTLYLTPTNASLFDIQEYGVYLAHQVQVPLQRLRIV
RTQTFEDNYALVNDGPNLNTTPTVTSAGSPGRLQLRLSLTEILKSGESDQILNRQ
LCYQTLTWKDIHFKNQNLALTIDITNRSRACPGCPMCKSGRCESEDQCLSTRT
VCAGCARCKGPLPTDCHECAAGCTGPKHSDCLACLFHNSGICELCHPALVYNT
DTFESMPNPEGRYTFGASCVTACPYNLSTDVSGCTLYCPHNOEVTADGTCRCEK
SKPCARVCYGLGMEHLEVRVATSANIQEACKIFGSLAFILPFSDFGPASNTAPL
QPEQLQVETLEEITGLYLYISAMPDSDLVSFQNLQVIRGRILHNGAYSLTLQGLGI
SWIGLSRELGLALIHNTFLCVHTVPDQLFRNPQALLTANRPEDECVEG
LACHOLCARGHCWPGQCVNCSQFLRQECVECRVQGLPREYVNAHCLPGHPE
COPONGSVTCFGEADQVACAHKDPFPCVACRCPGVKQDLSYMP IWKFPDEGACQ
PCPINCHSCVDLDDKGGCPAEQASPLSLISAVVGLLVVGLVGVGILIKRROKI
RKYTRRRLLQETELVEPLTPSGAMPNQAMRIILKTELKRVKVLGSGAGTGVKGIWI
PDGENKIPVAIKVIRENTSPKANEIIDEAYVMAGVSPVSRLLGICLTSTVOLVT
QLMPYCCLLDHYRENRLGSDLLNMCQIAGMSYLEDVRLVHRDLAARNVLKSP
NHVKTDFGLARLLDIDEYHADGKVPIKMALESILRRRTHQSDVMSYGVTVWE
LMTFGAKPYDGIPIAREIPDLLEKGRLPPOPICTIDVYMIWKCMIDSECRPREL
VSEFSMRARDPQRFVVIQNEDLGPASPLDSTFYRSLEDDEDDGLVDAYEYLPQGGF
FCPDPAAGAGVHHRRSSRSGGDLTLGLEPSEEAAPRSLAPSGAGSDVFDG
DIGMGAAGLQSLPHTDPSLPQYSEDPTVPLPSETDGVYVAPLTCSPQEVNPDVR
PQPPSPREGPLPAAPAGATLERPKTLSPGKNGVVKDVFAGVGAVENPEYLPQGGAA
PQHPPPAFSPAFDNLVYWDQPPERGAAPPSTFKGTPTAENPEYLGLDVPV"
175..237
/note="put. signal peptide (aa -21 to -1)"
mat_peptide 238..3939
/product="put. c-erb-B-2 protein (aa 1-1234)"
misc_feature 376..384
/note="pot. glycosylation site"
misc_feature 544..558
/note="pot. glycosylation site"
misc_feature 733..741
/note="pot. glycosylation site"
misc_feature 949..957
/note="pot. glycosylation site"
misc_feature 1762..1770
/note="pot. glycosylation site"
misc_feature 1885..1893
/note="pot. glycosylation site"
sig_peptide
mat_peptide
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

```

```

misc_feature 2059..2067
/note="pot. glycosylation site"
misc_feature 2353..3132
/note="aa 727-986, seq. homologous to EGF receptor kinase
domain"
misc_feature 2445..2454
/note="pot. glycosylation site"
misc_feature 4455..4460
/note="put. polyA signal"
polyA_site 4473
/note="polyA site"

BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN

Alignment Scores:
Pred. No.: 6.37e-187 Length: 4473
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 9 Gaps: 1

SEQ6 (1-919) x HSERB2R (1-4473)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGluYala 20
Db 175 ATGGAGCTGGCGGCTTGTGCGGCTGGGGCTCCTCTCGCCCTCTTGGCCCGGAGCC 234
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCCAAAGTGTCCACCGGCACACATGAGCTGGGCTGCCCTGCCAGTCCCGAG 294
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGGTGGCGAGGAACCTG 354
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 355 GAATCACTACCTGCGCCACCAATGCCAGCTGTCCTTCTGCGAGGATATCCAGGAGGTG 414
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGGCTACGTGCTCATCGCTCAACCAAGTAGGAGGCTGCCACTGCCAGGCTCCGG 474
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTGTGGAGGCCACCCAGCTCTTTGAGGACAATATGCCCTGGCCGTGCTAGACATGGA 534
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGAACATACACCCCTGTTCACAGGGGCTTCCCAGAGGCTCCCGGGAGCTG 594
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTCGAAGCCTCAAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG 654
Qy 161 LeuCystTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 655 CTCTGCTACCGAGACAGATTTGTGGAAGACATCTTCCACAGAACACACCACTGGCT 714
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAAG 774
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 775 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTCTCAGAGCCTCAGCGCACTGCTGT 834
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db 835 GCGGGTGGCTGTGCCCTCTCAAGGGGCCACTGCCCACTGACTGCTCCCATGAGCAGTGT 894
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260

```

```

Db 895 GCTCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCAACGAC 954
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 955 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGCTCACCTACACACAGACACACTTTGAG 1014
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 1015 TCCATGCCCAATCCCGAGGGCCGATATACATTCGCGCCAGCTGTGTGACTGCTGCTCC 1074
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1075 TACAACATACCTTTCTACGGACGTGGGATCTCGACCTCTCGTCTGCCCTCTGCACAAACAA 1134
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1135 GAGGTGACAGCAGAGGATGAACACACGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGGA 1194
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1195 GTGTGCTATGCTTGGGCATGGAGCACTTGGAGAGGTGAGGGCAGTTACCAAGTCCCAAT 1254
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1255 ATCCAGCAGTTTGTGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC 1314
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
Db 1315 TTTGATGGGACCCAGCTCCAACTGCCCGCTCCAGCAGCAGAGCTCCAAGTGT 1374
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1375 GAGACTCTGGAAGAGATCAGAGTTACCTATACATCTCAGCATGCGCGGACAGCCTGCCT 1434
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1435 GACCTCAGCTCTTCCAGAACCTGCAAGTAATCCGGGAGCGAATCTGCACAAATGGCGCC 1494
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1495 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCTACTCAGGGAA 1554
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1555 CTGGCAGTGGACTGGCGCTCATCCACATAACACCCACCTCTGCTTCGTGCACACGGTG 1614
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1615 CCCTGGACCAAGCTCTTTCCGAAACCCCGCACCAAGCTCTGTCTCCACACTGCCAACCGGCA 1674
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1675 GAGGACAGTGTGTGGCGAGGGCTGGCTGCCACAGCTGTGCGCCCGGAGGCACTGC 1734
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1735 TGGGTCCAGGGCCACCCAGTGTGTCACTGACGACAGTTCTTCTCGGGGCCAGGAGTGC 1794
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1795 GTGGAGGAATGCCAGTACTGCAAGGGCTCCCGAGGGCTCCCGAGGAGTATGTGAATGCCAGCACTGT 1854
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1855 TTGCCGTGCCACCTGAGTGTACGCCCCAGAAATGGCTCAGTGTGCTGTTTGGACCGGAG 1914
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1915 GCTGACCAAGTGTGGCTGTGCCACTATAGGACCTCCCTTCTGCTGCGGCGCGCTGC 1974
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620

```

Db 1975 CCACGGGTGTAACCTGACCTCTCTATATGCCCCATCTGGAAGTTTCCAGATGAGGAG 2034

QY 621 GlyAlaCysGlnProCysProIleasnCysThrHisSerCysValAspLeuAspAspLys 640

Db 2035 GGGCATGCGAGCTTGGCCCATCACTACCCACCTCTGTGTGACCTGGATGACAAAG 2094

QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653

Db 2095 GGCTGCCCGCGAGCAGAGAGCCAGCCCTCTGACGCTCATCTCTCGCGTGTGGC 2154

QY 653 ----- 653

Db 2155 ATTCTGCTGGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGAGCGAGCAG 2214

QY 653 ----- 653

Db 2215 AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGTGTGAGCGCGTG 2274

QY 653 ----- 653

Db 2275 ACACCTAGCGGAGCGATGCCAACCAAGCGGAGATGCGGATCCTGAAAGAGACGAGCTG 2334

QY 653 ----- 653

Db 2335 AGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCAGAGTCTCAAGGGCATCTGGATC 2394

QY 653 ----- 653

Db 2395 CCTGATGGGGAGATGTGAAATTCAGTGGCCATCAAGTGTGTGAGGAAACACATCC 2454

QY 653 ----- 653

Db 2455 CCCAAAGCAACAAGAAATCTTAGACGAAGCATACGTGATGCTGGTGTGGCTGCCCA 2514

QY 653 ----- 653

Db 2515 TATGTCTCCGCTTCTTGGGCATCTGCCTGACATCCACGGTGACAGCTT 2574

QY 653 ----- 653

Db 2575 ATGCCCTATGGTCTCTTAGACCATGTCCGGGAAACCGGAGCGCTGGCTGCCAG 2634

QY 653 ----- 653

Db 2635 GACCTGCTCACTGGTGTATGCAGATTGCCAAGGGATGAGTACCTGGAGGATGTGGG 2694

QY 653 ----- 653

Db 2695 CTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA 2754

QY 653 ----- 653

Db 2755 ATTACAGACTTCGGGTGCTCGGCTGTGGAGATTGACGAGACAGATACCATGCAGAT 2814

QY 653 ----- 653

Db 2815 GGGGCAAGGTGCCCATCAAGTGGATGGGCTGGAGTCCATTCTCGCGCGGCTTCACC 2874

QY 653 ----- 653

Db 2875 CACCAGAGTATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2934

QY 653 ----- 653

Db 2935 AAACCTTACATGGATGCCAGCCCGGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2994

QY 653 ----- 653

Db 2995 CTGCCCAAGCCCGCATCTGCACCATGTGATGCTACATGATCATGTGCAATGTTGGATG 3054

QY 653 ----- 653

Db 3055 ATTGACTCTGAATGTCCGCCAAGATTCGGGAGTTGGTGTCTGAATTTCTCCGCGATGGCC 3114

QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664

Db 3115 AGGGACCCCGCAGCGTTTGTGTATCCAGATAGGACTTGGGGCCAGCCAGTCCCTTG 3174

QY 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684

Db 3175 GACAGCACCTTCTACCGCTCACTGTGGAGGACGATGACATGGGGACCTGTGTGATGCT 3234

QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704

Db 3235 GAGGAGTATCTGTACCCAGCAGGCTTCTTCTGTCCAGACCTGCCCGCGCGTGGG 3294

QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724

Db 3295 GGCATGGTCCACAGCCAGCCAGCTCTCTACAGGAGTGGGGTGGGACCTGACA 3354

QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744

Db 3355 CTAGGCTGGAGCCCTCTGAAGAGGAGGCCCGCAGGTCTCCACTGGCACCTCCGAAGG 3414

QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyGlyLeuGlnSer 764

Db 3415 GCTGCTCCGATGTATTTGATGGTACCTTGGGAATGGGGGAGCCCAAGGGGCTCAAGG 3474

QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspPrcThrValProLeu 784

Db 3475 CTCCCCACATGACCCCGCAGCCCTCTACAGCGGTACAGTGGAGACCCACAGTACCCTG 3534

QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804

Db 3535 CCCTCTGAGCTGTGGTGTACCTTGCCTCCCTGACCTGCAGCCCGCAGCTGAATATGTG 3594

QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824

Db 3595 AACCCAGCAGATGTTCGGCCCGCAGCCCTTTCGCCCGGAGAGGGCCCTCTGCTGCTGCC 3654

QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844

Db 3655 CGACCTGTGTGGTCCACTCTGGAAGGCCACAGACTCTCTCCCGAGGAGATGGGCTC 3714

QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864

Db 3715 GTCAAAGACGTTTTTGCCTTTGGGGTGCCTGGAGAACCCCGAGTACTGTACACCCAG 3774

QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884

Db 3775 GGAGGAGTGGCCCTTCAGCCCGCAGCTCTCTGCTTCAGCCCGAGCTTCGACACCTC 3834

QY 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904

Db 3835 TATTACTGGGACGAGCCAGGAGGGGGGCTCCACCCAGCAGCCCTTCAAGAGGACA 3894

QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919

Db 3895 CCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGAGCTGCCAGTG 3939

RESULT 9

AX060704

LOCUS

DEFINITION

AX060704

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE

AUTHORS

TITLE

JOURNAL

AX060704 3768 bp DNA linear PAT 22-JAN-2001

Sequence 2 from Patent WO0100244.

AX060704.1 GI:12406101

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Erickson, S. and Schwall, R.

Methods of treatment using anti-erbB antibody-maytansinoid conjugates

Patent: WO 0100244-A 2 04-JAN-2001;

Genentech, Inc. (US)

Location/Qualifiers

1. 3768

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 758 a 1170 c 1121 g 719 t

ORIGIN

Alignment Scores:

Pred. No.: 1.12e-186 Length: 3768

Score: 4892.00 Matches: 918

Percent Similarity: 73.15% Conservative: 0

Best Local Similarity: 73.15% Mismatches: 1

Query Match: 96.34% Indels: 336

DB: 6 Gaps: 1

SEQ6 (1-919) x AX060704 (1-3768)

Qy 1 MetGluLeuAlaAlaLeuCysArgTrrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 Db 1 ATGGAGCTGGCGGCTTGTGGCGTGGGGCTCTCTCGCCCTTGGCCCGGGAGCC 60
 Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 61 GCGAGCACCAAGTGTGACCGGCACACATGAAGCTGCGGCTCCCTGCGCAGTCCCGAG 120
 Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnLeu 60
 Db 121 ACCCACTGGAGATGCTCCGCGACCTCTTACCAGGGCTGCCAGGTGGTGACGGGAACCTG 180
 Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 181 GAATCACTACCTGCGCCACCAATGCCAGCTGTCTCTCTGAGGATATCCAGAGGTG 240
 Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 241 CAGGGCTACGTGCTCATCGCTCACAACCAAGTAGGCGAGTCCCACTGCAGAGGCTCGG 300
 Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 301 ATTGTGGAGGACCCAGCTCTTGGAGACAACTATGCCCCGTGGCGGTGCTAGACAATGGA 360
 Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 361 GACCGCTGAACATACCAACCCCTGTACAGGGGCTCCCGAGAGGCTCGGGAGCTG 420
 Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
 Db 421 CAGCTTCAGACCTCAGAGATCTTGAAGAGGGGTCTTGTATCCAGCGGAACCCCGAG 480
 Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 Db 481 CTCGTACCAAGGACAGATTTGTGAAGGACATCTTCCACAAGAACCAACAGCTGGCT 540
 Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTTGCACCCCTGTCTCCGATGTGAAG 600
 Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 601 GGCTCCCGCTGCTGGGGAGAGATTTGTGAGGATTTGTGAGAGCTTCAGAGCTTGAACGCGCACTGTCTGT 660
 Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
 Db 661 GCCGTGGCTGTGGCGCTGCAAGGGGCACTGCCCACTGACTGCTGCGCATGAGCAGTGT 720
 Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 721 GCTGCGGCTGCACGGGGCCCAAGACACTCTACATGCGCTGGCTGCTCCCTCCACTTCAACCCAC 780
 Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 1861 GGCACATGCCAGCCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1920

Db 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGTGTACCTACAAACACACACACGTTTGAG 840
 Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 841 TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTACCTGCTGCTGCC 900
 Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 901 TACAACCTACCTTTCTACGGACGTGGATCCTGACCCCTCGTCTGCCCTTCGCAACCA 960
 Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTCAGCAAGCCCTGTGCCGA 1020
 Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1021 GTGTGCTATGCTTGGGCGATGGAGCTTGGAGAGGTGAGGGCAGTTTACCAGTGCCTAAT 1080
 Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 Db 1081 ATCCAGGAGTTTGTGGCTGCAAGAGATCTTGGGAGCCTGGCATTTCTGCCGAGAGC 1140
 Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 Db 1141 TTTGATGGGACCCAGGCTCCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1200
 Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrrpAspSerLeuPro 420
 Db 1201 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGGCCGGCAGCGCTGCT 1260
 Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1261 GACCTCAGGCTTCCAGAACCTGCAAGTATATCCGGGACCAATCTCCACAAATGGCGCC 1320
 Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1321 TACTCGTGCACCTGCAAGGCTGGGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1381 CTGGGAGTGGACTTGGCCCTCATCCATTAACACCCACCTCTGCTGCTGCTGCTGCTGCTG 1440
 Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1441 CCCTGGGACCAAGCTTTCGGAACCCGACCAAGCTGCTGCTCCACACTGCCAAGCGGCA 1500
 Qy 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 Db 1501 GAGCAGAGTGTGTGGGCGAGGGCTTGGCTTGCACCAAGCTGCTGCTGCTGCTGCTGCTG 1560
 Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 Db 1561 TGGGCTCAGGGCCACCCAGTGTCAACTGACAGCAGATTCCTTCGGGGCCAGAGTGC 1620
 Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 Db 1621 GTGGAGCAATCCGAGTGTCCAGGGCTCCCGAGGAGTATGTGAATGCCAGCAGCTGT 1680
 Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1681 TTGCGGTGCCACCTGAGTGTGACGCCAGAAATGGCTCAGTGTGCTGCTGCTGCTGCTGCTG 1740
 Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
 Db 1741 GCTACCAAGTGTGGGCTGTGGCCACTATGAAGACCTTCCCTTCTGCTGCTGCTGCTGCTG 1800
 Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
 Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 Db 1861 GGCACATGCCAGCCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1920

QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer-----	653	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	1921	GGCTGCCCGCGGAGCAGAGACCAGCCCTCTGACGTCCATCGTCTCTGGGGTGTGGC	1980	3001	GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGGTGGATGCT	3060
QY	653	-----	653	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	1981	ATTCTGCTGGTGGTGGTCTTGGGGGTGTCTTTGGGATCCTTCATCAAGCGACGGCAGCAG	2040	3061	GAGGAGTATCTGGTATCCCAACAGCAGGGCTTCTTCTGTCCAGACCTGCCCGGGCGCTGGG	3120
QY	653	-----	653	705	GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr	724
Db	2041	AAGATCCGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTG	2100	3121	GGCATGGTCCACACAGCAGCACCGCAGCTCATCTACCAGAGTGGCGGTGGGACCTGACA	3180
QY	653	-----	653	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	2101	ACACCTAGCGGAGCGATGCCAACACGAGCGCAGATGCGGATCCTGAAAGAGACGGAGCTG	2160	3181	CTAGGCGTGGAGCCCTCTGAAGAGGAGGAGGCCCCAGGTCTCCACTGGCACCTCCGAAGGG	3240
QY	653	-----	653	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
Db	2161	AGGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTTACAAAGGCATCTGGATC	2220	3241	GCTGGCTCCGATGTATTGTATGGTGACCTGGGAATGGGGGAGCCCAAGGGCTGCAAGC	3300
QY	653	-----	653	765	LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu	784
Db	2221	CCTGATGGGAGAAATGTGAATAATCCAGTGGCCATCAAAGTGTGAGSAAACACATCC	2280	3301	CTCCCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAGGAGCCCCACAGTACCCCTG	3360
QY	653	-----	653	785	ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal	804
Db	2281	CCCAAGCCCAACAAGAAATCTTAGACGAAGCATACGTATGCTGTTGGGTCCCA	2340	3361	CCCTCTGAGACTGATGGTACGTTGCCCTGACCTGCAGCCGCCAGCCTGAATATGTG	3420
QY	653	-----	653	805	AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla	824
Db	2341	TATGTCCTCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400	3421	AACCAAGCAGATGTTGGCCCCAGCCCCCTTCGCCCGAGAGGGCCCTCTGCCTGTGCTGC	3480
QY	653	-----	653	825	ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal	844
Db	2401	ATGCCCTATGGTGCCTCTTAGACCATGTCCGGGAAACCGCGACGCTGGGCTCCAG	2460	3481	CGACCTGCTGGTGCCACTCTCGAAAGGGCCAGACTCTCTCCCAGGGAAGATGGGGTC	3540
QY	653	-----	653	845	ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln	864
Db	2461	GACCTGCTGAAGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGCGG	2520	3541	GTCAAAGACGTTTTTGGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGCACACCCAG	3600
QY	653	-----	653	865	GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu	884
Db	2521	CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGTTCAAGAGTCCCAACCATGTCAA	2580	3601	GGAGAGCTGCCCCCTCAGCCCCACCTCCTCCTGCTTCAGCCCGAGCCTTCGACAAACCTC	3660
QY	653	-----	653	885	TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr	904
Db	2581	ATTACAGACTTCGGGCTGGCTGGCTGGACATTGACGACAGAGTACCATGCAGAT	2640	3661	TATTACTGGACACGAGCCACCAGAGCGGGGGGCTCCACCCAGCACCTTCAAGAGGACA	3720
QY	653	-----	653	905	ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal	919
Db	2641	GGGGCAAGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCCTCCGCGCGGTTTCACC	2700	3721	CCTACGGCAGAAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG	3765
QY	653	-----	653	RESULT 10		
Db	2701	CACCAGAGTATGTGGAGTTATGGTGTACTGTGTGGAGCTGATGACTTTTGGGGCC	2760	AX467229	Sequence 3 from Patent WO0234287.	linear
QY	653	-----	653	LOCUS		
Db	2761	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGTGAAAGGGGAGCGG	2820	AX467229	Accession	
QY	653	-----	653	DEFINITION		
Db	2821	CTGCCCCAGCCCCCATCTGCACCAATTGATGTCTACATGATCATGTGTTGATG	2880	AX467229	Version	
QY	653	-----	653	KEYWORDS		
Db	2881	ATTGACTCTGAATGTCGGCCAAAGATTCGGGAGTTGGTGTCTGAATTCCTCCCGCATGGCC	2940	AX467229.1	GI:21900511	
QY	654	-----	654	SOURCE		
Db	2941	AGGACCCCCAGCGCTTTGTGGTCTATCCAGAATGAGGACTTGGGCCCGCCAGCTCCCTTG	3000	ORGANISM		
				Homo sapiens (human)		
				Homo sapiens		
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
				Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
				1		
				Beier, A.M., Gautam, A. and Mouritsen, S.R.		
				Novel therapeutic vaccine formulations		
				Patent: WO 0234287-A 3 02-MAY-2002;		
				Pharmexa A/S (DK)		
				Location/Qualifiers		
				1..3768		
				/organism="Homo sapiens"		
				/mol_type="genomic DNA"		
				/db_xref="taxon:9606"		

CDS	1..3768	Db	481	CTCTGCTACGAGCACGATTTTGTGGAAGGACATCTTCCACAAGAACCAACGCTGGCT	540
/note="unnamed protein product"	/codon_start=1	Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
/protein_id="CAD42452.1"	/db_xref="GI:21900512"	Db	541	CTCACACTGATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTTTCCGATGTGAAG	600
/translation="MELAAALRWGLLALLPPGAASTQVCTGTDMLRLPASPETHLD		Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
RGTHLYGCQVQGNLELYLPTNASLFLQDIOEVQGYVLIAHNQVQVPLQRLRV		Db	601	GGCTCCCTCTCTGGGAGAGAGTCTCAGGATTTGACAGCCTCAGCGGCACTCTCTGT	660
LRGLQFEDNYALVDNMGDPLNNTTPVTGASPGGLRELQLRSLTEILKGGVLIQRNP		Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys	240
LCYDITLWKDIPKHNQALTLITNRSRACHPCSPMKGSRGWCWGESSEDCOSLITRT		Db	661	GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCACTGACTGTGCTCATGAGCAGTGT	720
VCAGTCARCGPLPTDCCHEQCAAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNT		Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
DTFESMNPGRYTFGASCVTACTPYNILSTDVSGCTLVCPLNHNEVTAEDDQRCCK		Db	721	GCTCCGGCTGCACGGGGCCCCAAGCACTCTGACTGCTGGGCTGGCTCCACTTCAACCA	780
SKPCARVCYGLGMEHLREVRVTSANTIOEFAGCKKIFGSLAFLPESDGDPSANTAPL		Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
OPEQLQVFTELEEDTGYLYTSAPDLSPLDSVFQNLQVIRGILHNGAYSLTLQGLGI		Db	781	AGTGCATCTGTGAGCTGCACCTGCCACCCCTGGTGCCTACCAACACAGACACGTTT	840
SWLGLRELGLGIALITHNTHLCFVHTVPDOLFRNPHOALLHTANRPDECVGEG		Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
LACHOLCARHCWCPGPTQCVCNCSQFLRGQECVECRVLQGLPREYVNAHCLPCHPE		Db	841	TCCATGCCCAATCCGAGGGCGGTATACATTCCGGCGCAGCTGTGTGACTGCTGTCCC	900
COPONGSWCTGPEADOCVACHYKDPPFCVARGCPGKPDLSYMPNWKFPDEBGAQ		Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
PCPYNTHSCVDLDDKCPAQRASPLTISVAVVGLLVVGLVVGFGILIKRRQKI		Db	901	TACAACCTACCTTTCTACGGAGCTGGGATCTCTGCACCCCTCGTGTGCCCTGCACAA	960
PDGENVKIPVAIKVIRENTSPKANKEILDEAYVMAGVSPYVSRLLGICLTSTVOLVT		Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
OLMPYGLLDHVRNRLGTSQDILLNMCQIAGMSYLEDLVRLVHRLDLAARNVLKSP		Db	961	GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAACTGCAGCAACCCCTGTG	1020
NHVKITDFGLARLLDIDETEHADGKVPKIMMALESILRRRFTHQSDVMSYGVTVWE		Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
LMTGAKPYDGIIPARETIDLEKGERLPDPPICTIDVIMVIMKCMWIDSECRPFREL		Db	1021	GTGTGCTATGCTTGGGCATGGACACTTGCAGAGGTGAGGGCAGTTTACCAGTGCCAAT	1080
VSESRMARDPORFVIONEDLGPASPLDSFYRSLLEDDDMGDLVDAEYLVPQQGF		Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
FCPDPAAGAGSNVHHRSSSTRSGGDLTLGLPSEEARPSPLASEGAGSDVDFG		Db	1081	ATCCAGGAGTTTGTGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTTCTGCCGGAG	1140
DLGMAAGKGLQSLTHDPSPQLQRYSEDTVPLPSETDGYVAPLTCSPQPEYVNPQDVR		Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
PQPSREGPLPAAPAGATLERAKTILSPGKNGYVKDVFAGGAVENPEYLTPOGGAA		Db	1141	TTTGTATGGGACCCAGCCTCCACACTGCCCGCTCCAGCCAGCAGACGCTCCAAGT	1200
PQPHPPAFSPAFNLVYWDODPERGAPPSTFKGTPTAENPEYLGLDVpv"		Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro	420
BASE COUNT 758 a 1170 c 1121 g 719 t		Db	1201	GAGACTCTGGAAGAGATCACAGGTATACCTATACATCTCAGCATGCGCGACAGCTG	1260
ORIGIN		Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Alignment Scores:		Db	1261	GACCTCAGCGCTTCCAGAACCTCAAGTAATCCGGGGACGAAATTTCTGCAATGGCG	1320
Pred. No.: 1.12e-186	Length: 3768	Qy	441	TyrSerLeuThrLeuGlnGlyIleSerTyrLeuGlyLeuArgSerLeuArgGlu	460
Score: 4892.00	Matches: 918	Db	1321	TACTCGTGACCCCTGCARAGGCTGGGCATCAGCTGGCTGGGCTGGCTCTACTGAG	1380
Percent Similarity: 73.15%	Conservative: 0	Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Best Local Similarity: 73.15%	Mismatches: 1	Db	1381	CTGGCAGTGAGCTGGCCCTCATCCACCATTAACACCCACCTCTGCTGCTGCACAG	1440
Query Match: 96.34%	Indels: 336	Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
DB: 6	Gaps: 1	Db	1441	CCCTGGGACCAAGCTCTTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCA	1500
SEQ6 (1-919) x AX467229 (1-3768)		Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Qy 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla		Db	1501	GAGGACGAGTGTGGGCGAGGGCTGGCTGCCACCAAGCTGTGCGCGCCGAGGCACTGC	1560
Db 1 ATGAGAGCTGGCGGCTTGTGCGCTGGGGGTCTCTCTCGCCCTTGTGCCCGGAG		Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu		Db			
Db 61 GCGAGACCCCAAGTGTGCACCGCACAGACATGAGCTGCGGCTCCCTGCCAGTCC		Qy			
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu		Db			
Db 121 ACCCACCTGGACATGCTCCGCGCACCTCTACACAGGCTGCCAGGTGGTGACG		Qy			
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal		Db			
Db 181 GAATCTACCTACCTGCCACCAATGCCAGCTGTCTCTTCTGAGGATATCCAGGAG		Qy			
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg		Db			
Db 241 CAGGGCTACGTGCTCATCGCTCACACCAAGTAGGAGGAGTCCCACTGCAGAGG		Qy			
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly		Db			
Db 301 ATTTGCGGAGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGCTAGACA		Qy			
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu		Db			
Db 361 GACCGCTGGAACAATACCAACCCCTGTACAGGGGCTCCCGCAGGAGGCTGCGG		Qy			
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln		Db			
Db 421 CAGCTTGAAGCCCTCACAGAGATCTTGAAGAGGAGGGTCTTTGATCCAGCGGA		Qy			
Qy 161 LeucCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla		Db			

Db	1561	TGGGGTCCAGGGCCCCCAGTCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC	1620	QY	653	-----	653
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560	Db	2701	CACCAGATGATGTGGAGTTATGTGTCACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
Db	1621	GTGGAGATGCCAGTACTGCAGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1680	QY	653	-----	653
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580	Db	2761	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG	2820
Db	1681	TTGCCGTGCCACCTCAGTGTACAGCCCAAGATGGCTCAGTACCTGTTTGGACCGAG	1740	QY	653	-----	653
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600	Db	2821	CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTCAAATGTTGGATG	2880
Db	1741	GCTGACCACTGTGGCGCTGTGCCACTATAGACCCCTCCCTTCTGCGTGGCGCGCTGC	1800	QY	653	-----	653
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620	Db	2881	ATTGACTCTGAATGTCCGCCAAGATTCCTGGGAGTTGGTGTCTGAANTTCTCCCATGCGCC	2940
Db	1801	CCAGCGGTGTGAACCTGACCTCTCTACATCCCATCTGGAAGTTTCCAGATGAGGAG	1860	QY	654	-----	654
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640	Db	2941	AGGGACCCCGAGCGCTTTGTGGTCAATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTG	3000
Db	1861	GGCGCATGCCAGCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAG	1920	QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653	Db	3001	GACAGCACCTTCTACCGCTCACTGTGGAGGACATGACATGGGGACCTGTGTGATGCT	3060
Db	1921	GGCTGCCCGCCGAGCAGAGACAGCCCTCTGACGTCCATCTCTGCGGTGGTTGGC	1980	QY	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
QY	653	-----	653	Db	3061	GAGGATATCTGTATCCCGAGCGGCTTCTTCTGTCAGAGACCTGCCCGGCGCTGGG	3120
Db	1981	ATTCTGCTGCTGTGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAG	2040	QY	705	GlyMetValHisHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr	724
QY	653	-----	653	Db	3121	GGCATGGTCCACCACAGCGCACCGACCTCATCTACCAAGGAGTGGCGTGGGGACCTTGACA	3180
Db	2041	AAGATCCGGAAGTACAGATGCGGAGACTGCTGCAGAAACGGAGCTGGTGAGCGCGCTG	2100	QY	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
QY	653	-----	653	Db	3181	CTAGGCTGGAGCCCTCTGAAGAGGAGGCCGCCAGGTCTCCACTGGCACCTCCGAGGG	3240
Db	2101	ACACCTAGCGGAGCGATGCCACACAGCGGAGATCGGATCCTGAAAGAGACGGAGCTG	2160	QY	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
QY	653	-----	653	Db	3241	GCTGGCTCGATGATTTGATGTGACCTGGGAAATGGGGCAGCCCAAGGGGCTGCAAGC	3300
Db	2161	AGGAAGGTGAAGTGTGGNYCTGGCTGGCTTTTGGCACAGTCTACAAAGGCATCTGGATC	2220	QY	765	LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu	784
QY	653	-----	653	Db	3301	CTCCACACATGACCCACCCCTCTACAGCGGTACAGTGAGGACCCACAGTACCCCTG	3360
Db	2221	CCTGATGGGAGAAATGTGAAATTTCCAGTGGCCATCAAAGTGTGAGGAAACACATCC	2280	QY	785	ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal	804
QY	653	-----	653	Db	3361	CCCTCTGAGACTGATGGCTACGTTGCCCCCTTGACCTGCAGCCCCCAGCTGAATATGTG	3420
Db	2281	CCCAAAAGCCAAAGAAATCTTTAGACGAAGCATACGTGATGGCTGGTGTGGGCTCCCCA	2340	QY	805	AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla	824
QY	653	-----	653	Db	3421	AACCAACCAAGATGTTGGCCCCCAGCCCTTCCGCCCGAGAGGGCCCTCTGCTGCTGCC	3480
Db	2341	TATGTCTCCGCCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400	QY	825	ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal	844
QY	653	-----	653	Db	3481	CGACCTGTGGTGGCTCTGGAAGGGCCCAAGACTCTCTCCCGCAGGAGAAATGGGTC	3540
Db	2401	ATGCCCTATGGCTGCCTTTAGACCATGTCGGGAAACCGCGGACGCTGGGCTCCCGAG	2460	QY	845	ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln	864
QY	653	-----	653	Db	3541	GTCAAAGAGCTTTTTCCTTTGGGGTGGCGTGAGAAACCCCGAGTGTGACACCCAG	3600
Db	2461	GACCTGCTGAACCTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2520	QY	865	GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu	884
QY	653	-----	653	Db	3601	GGAGGAGCTGCCCTCAGCCCCCACCCTCTCTCTCCCTTCCAGCCAGCTTCGACACCTC	3660
Db	2521	CTCGTACACAGGGACTTGGCCGCTCGAAGAGTGTGGTCAAGAGTCCCAACCATGTCAA	2580	QY	885	TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr	904
QY	653	-----	653	Db	3661	TATTACTGGACCAAGACCCACAGAGCGGGGCTCCACCCAGCACCTTCAAAGGACA	3720
Db	2581	ATTACAGACTTCGGGCTGGCTGGCTGCTGGACATTGACGAGACAGATACCATGAGAT	2640	QY	905	ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal	919
QY	653	-----	653	Db	3721	CTTACGGCAGAGAACCAGAGTACCTGGGTCTGGAGCTGCCAGTG	3765

```

RESULT 11
AX481438          AX481438          3768 bp      DNA      linear      PAT 16-AUG-2002
LOCUS
DEFINITION      Sequence 52 from Patent WO02055693.
ACCESSION      AX481438
VERSION        AX481438.1  GI:22316352
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Kreutzer,R., Limmer,S., Rost,S. and Hadwiger,P.
TITLE        Method for inhibiting the expression of a target gene
JOURNAL      Patent: WO 02055693-A 52 18-JUL-2002;
Ribopharma AG (DE)
FEATURES
source
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT      758 a 1170 c 1121 g 719 t
ORIGIN
Alignment Scores:
Pred. No.:      1.12e-186      Length:      3768
Score:          4892.00      Matches:      918
Percent Similarity: 73.15%      Conservative: 0
Best Local Similarity: 73.15%      Mismatches: 1
Query Match:    96.34%      Indels:      336
DB:            6      Gaps:      1
SEQ6 (1-919) x AX481438 (1-3768)
Qy      1 MetGluLeuAlaLaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProProGlyAla 20
Db      1 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCGCCCTCTTGGCCCGGAGCC 60
Qy      21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      61 GCGAGCACCCNAGTGTGCACCGGCACACATGAGCTGGCGCTCCCTGCGCAGTCCCGAG 120
Qy      41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db      121 ACCCACCTGGACATGCTCGCCACCTCTACCGAGGCTGCCAGGTGGTGACGGGAACCTG 180
Qy      61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db      181 GAACCTACCTACCTGCCCGACCAATGCCAGCGCTGTCTTCCTGCGAGGATATCCAGGAGGTG 240
Qy      81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      241 CAGGGCTACGTGCTCATCGCTCACAACTAGTGGGAGGCTGCCACCTGCGAGGCTCGGG 300
Qy      101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      301 ATTGTGGAGCACCCAGCTCTTTGAGACAACTATGCCCTGGCGTGGCTAGACAAATGA 360
Qy      121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      361 GACCCGTGAAACATACACCCCTGTTCACAGGGGCTTCCACAGGAGGCTTCCCGAGGAGCTG 420
Qy      141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db      421 CAGCTTCGAAGCCCTCACAGAGATCTTGAAGAGGAGGCTTTGATCCAGCGGAACCCCGAG 480
Qy      161 LeuCystTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      481 CTCTGCTACCGAGCACGATTTTGTGAAGACATCTTCCACAGACACACACAGCTGGCT 540
Qy      181 LeuThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      541 CTCACACTGATAGACACCAACCGCTCTCGGGCTTCCAGGCTGCCACCCCTGTCTTCTCCGATGTGTAAG 600
201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      601 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGAGATTGTTCAGAGCCTTCAGAGCAGCTGTCTGT 660
Qy      221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db      661 GCCGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCACTGACTGCTGCTGCCATGAGCAGTGT 720
Qy      241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      721 GCTCCCGGCTGCAGGGCCCCAAGCACTCTGACTGCTGGCTGCTCCCTCCACTTCAACCCAC 780
Qy      261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGCTCACACACACAGACACGTTTGAG 840
Qy      281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      841 TCCATGCCCAATCCGAGGGCCGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTGCC 900
Qy      301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      901 TACAACCTACCTTCTACGGAGCTGGGATCTCTGCCCTCTGCTGCCCTCCACCAACCAA 960
Qy      321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      961 GAGGTGACAGCAGAGGATGGAACACACACGGGTGTGAGAAGTGCACAGACCCCTGTGCCGA 1020
Qy      341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1021 GTGTGCTATGCTGTGGCCTGGAGCACTTGGAGAGGTGGAGGCGAGTTACCAATGCCAAT 1080
Qy      361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      1081 ATCCAGAGATTGTGTGCTGCAAGAAGATCTTTGGAGCCTTGGCATTTCTGCCGGAGAGC 1140
Qy      381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      1141 TTTGATGGGACCCAGCCTCCACACCTGCCCGCTCCAGCCGAGAGAGCTCCAAGTGT 1200
Qy      401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db      1201 GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGSCCGGACAGCCTGCCT 1260
Qy      421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db      1261 GACCTCAGCGCTCTTCCAGAACCTGCAAGTAAATCCGGGGAGCAATCTGCACAAATGGGCC 1320
Qy      441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1321 TACTCGCTGACCTTGCAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGAA 1380
Qy      461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1381 CTGGCAGTGGACTGGCCCTCATCCACCAATACACCCACCTCTCTCTTCGTGCACACGGTG 1440
Qy      481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db      1441 CCCTGGACACAGCTCTTTCCGAACCCGACCAAGCTCTGCTGCCACACTGCCAACCGGCCA 1500
Qy      501 GluAspGluCysValGlyGluLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db      1501 GAGGACAGGTGTGTGGGCGAGGGCTGGCTGCCCTGCCAGCTGTGCGCGGAGGCACTGC 1560
Qy      521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db      1561 TGGGGTCCAGGGCCCCACCTGCTGCTCACTGACGACAGTTCCTTCTGGGGCCAGGAGTGC 1620
Qy      541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db      1621 GTGGAGGAATGCCGAGTACTGTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680

```

QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1681 TTGCGGTGCCACCTGAGTGTGAGCCCGAGATGCTCAGTACCTGTTTGGACGGAG 1740
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 Db 1741 GCTGACAGTGTGGCGTGTGCCCACTATAAGGACCTCCCTTCTGCGTGGCGGCTGC 1800
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 Db 1801 CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 Db 1861 GCGCATGCCAGCTTGCCCCCACTCACTGCACCCCACTCTCTGTGTGACCTGATGACAG 1920
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer- 653
 Db 1921 GGCTGCCCGCGAGCAGAGACCCCTCTGAGCTCCATCGTCTCGGTGGTTGGC 1980
 QY 653 ----- 653
 Db 1981 ATTCTGCTGGTGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGGAGCAG 2040
 QY 653 ----- 653
 Db 2041 AAGATCCGGAAGTACACGATGCGAGACTGCTGCAGAAACGGAGCTGTGTGGCGCGTG 2100
 QY 653 ----- 653
 Db 2101 ACACCTAGCGGAGCGATGCCAACCGCGCAGATGCGGATCCTGAAGAGACGAGGCTG 2160
 QY 653 ----- 653
 Db 2161 AGAAGGTGAAGTGTGGATCTGGCGTCTTTGGCACAGTCTACAAGGGCATCTGGATC 2220
 QY 653 ----- 653
 Db 2221 CCTGATGGGAGAAATGTGAATTCACGTGGCCATCAAAGTGTGAGGAAACACATCC 2280
 QY 653 ----- 653
 Db 2281 CCCAAGCCACAAGAAATCTTAGACGAAGCATACGTATGGTGGTGGGCTCCCA 2340
 QY 653 ----- 653
 Db 2341 TATGTCTCCGCTTCTGGGCATCTGCCCTGACATCCACGGTGCAGTGTGACACAGCTT 2400
 QY 653 ----- 653
 Db 2401 ATGCCCTATGGTGCCTCTAGACCATGTCCGGGAAACCGGCGCTGGGCTCCAG 2460
 QY 653 ----- 653
 Db 2461 GACCTCTGAAGTGTGTATGAGATGTCAGATGCCAAGGGGATGAGTACCTGAGGATGGCG 2520
 QY 653 ----- 653
 Db 2521 CTCGTACACAGGACTTGGCCGCTCGAAGCTGCTGGTCAAGAGTCCCAACCATGTCAAA 2580
 QY 653 ----- 653
 Db 2581 ATTACAGACTTGGGCTGGCTGGGTGCTGGACATTGACGAGACAGATACCATGCAGAT 2640
 QY 653 ----- 653
 Db 2641 GGGGGCAAGGTGCCCATCAAGTGTGGGTGGGTGGAGTCCATCTCGCGCGGGTTACCC 2700
 QY 653 ----- 653
 Db 2701 CACGAGATGATGTGGAGTTATGGTGTACTGTGGGAGCTGATGACTTTTGGGGCC 2760
 QY 653 ----- 653

Db 2761 AAACCTTACGATGGATCCACCGCGGAGATCCCTGACCTGCTGGAAGAGGGGAGCGG 2820
 QY 653 ----- 653
 Db 2821 CTGCCCCAGCCCCCATCTGCACCATGTGATGTCTATCATGATCATGTCACAAATGTTGGATG 2880
 QY 653 ----- 653
 Db 2881 ATTGACTCTGAATGTGCGGCCAAGATTCCGGGAGTTGGTCTCTGAATTCTCCCGCATGGCC 2940
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 Db 2941 AGGGACCCCCAGCGCTTTGTGTGTCATCCAGATAGGACTTGGGCCCAACCCAGTCCCTTG 3000
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 584
 Db 3001 GACAGACCTTACCGCTCACTGTGTGGAGACGATGACATGGGACCTGGTGGATGCT 3060
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
 Db 3061 GAGGAGTATCTGGTACCCAGCAGGGCTTCTTCTTCACAGCCCTGCCCGGGCGCTGGG 3120
 QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
 Db 3121 GGCATGGTCCACACAGCAGCCAGCTCATCTACAGAGTGGCGGTGGGACCTGACAC 3180
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 Db 3181 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCGCCAGGTCTCCACTGCGACCTCCGAAGGG 3240
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
 Db 3241 GCTGGCTCGATGATTTGATGGTACCTGGGAATGGGGCGACGCAAGGGCTGCCAAGC 3300
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 Db 3301 CTCCACACATAGCCCAACCCAGCTCTACAGCGGTACAGTGGAGACCCACAGTACCCCTG 3360
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 Db 3361 CCCTCTGAGACTGATGGCTACGTGGCCCTCTGACCTGCGAGCCCGCCAGCTGAATATGTG 3420
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
 Db 3421 AACCGACAGATGTTCCGGCCCGACCCCTTCGCCCGAGAGGGCCCTCTGCTGCTGCC 3480
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
 Db 3481 CGACCTGTCTGTGCCACTCTGGAAGGGCCAAAGACTCTCTCCCGAGGGAAGATGGGGTC 3540
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
 Db 3541 GTCAAGAGCTTTTGGCTTTGGGTGCGCGTGGAGAACCCGAGTACTTGACACCCAG 3600
 QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3601 GGAGGAGCTGCCCTCAGCCCGACCCCTCTCTGCTGCTTCAGCCAGCTTCGACAACTC 3660
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
 Db 3661 TATTACTGGGACGAGGACCCACAGAGCGGGGGCTCCACCCAGCAGCTTCAAGAGGACA 3720
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3721 CCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG 3765

RESULT 12
 AR202597
 LOCUS
 DEFINITION
 AR202597
 ACCESSION
 VERSION

AR202597 4530 bp DNA linear PAT 20-JUN-2002
 Sequence 4 from patent US 6365151.
 AR202597
 AR202597.1 GI:21498768

KEYWORDS

SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 4530)
 AUTHORS Halpern, M.S. and England, J.M.
 TITLE Cellular immunogens comprising cognate proto-oncogenes
 JOURNAL Patent: US 6365151-A 4 02-APR-2002;
 FEATURES
 Location/Qualifiers
 1..4530
 /organism="unknown"

BASE COUNT 922 a 1382 c 1346 g 880 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.34e-186 Length: 4530
 Score: 4892.00 Matches: 918
 Percent Similarity: 73.15% Conservative: 0
 Best Local Similarity: 73.15% Mismatches: 1
 Query Match: 96.34% Indels: 336
 DB: 6 Gaps: 1

SEQ6 (1-919) x AR202597 (1-4530)

Qy 1 MetGluLeuAlaLaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 Db 151 ATGGAGCTGGCGGCTTGTGCGGCTGGGGCTCCTCGCCCTTTGCCCGGAGCC 210
 Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 211 GCGAGCACCCAAAGTGTGCACGGGCACACATGAGCTGGGCTCCCTGCCAGTCCGCGAG 270
 Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlnAsnLeu 60
 Db 271 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAAACCTG 330
 Qy 61 GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 331 GAACCTACCTACCTGCCACCAATGCCAGCTGTCTTCTGACAGGATATCCAGGAGGTG 390
 Qy 81 GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 391 CAGGGCTACGGTGCATCGCTCACACCAAGTAGGAGGCTGCCACCTGCAGAGGCTCGG 450
 Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGly 120
 Db 451 ATTGTGGCAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAAATGGA 510
 Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 511 GACCCGTGAACAATACCAACCCCTGTACAGGGGCTCCCGCAGGAGGCTCGGAGCTG 570
 Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 Db 571 CAGCTTCAAGACCTTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 630
 Qy 161 LeuCysTyrrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 Db 631 CTCTGCTACCAAGACAGATTTGTGAAGGACATCTTCCACAGAACACACAGCTGGCT 690
 Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 691 CTCACATGATAGACACCAACCGCTCTCGGCGCTGCCACCCCTGTTCTCCGATGTGAAG 750
 Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 751 GGCTCCCGCTGCTGGGAGAGATTCTGAGGATTTGTGAGCCCTCAGACCTCGACGCTGCTGT 810
 Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
 Db 811 GCCGGTGGCTGTGCCCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 870
 Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260

Db 871 GCTGCCGCTGCACGGGCCCCAAGCACTCTGACTGCTGGCTCCCTCCACTTCAACAC 930
 Qy 261 SerGlyTleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGlu 280
 Db 931 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGTGTACCTACACACAGACACGCTTTGAG 990
 Qy 281 SerMetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 991 TCCATGCCAATCCGAGGCGGTATACATTCGCGCCAGCTGTGTGACTGCTGTCTCC 1050
 Qy 301 TyrAsnTyrrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 1051 TACAACATACCTTTCTACGGAGCTGGGATCTGTCCACCTCGTCTGCCCTGCACAACCAA 1110
 Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 1111 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGACAGCCCTGTGCCGA 1170
 Qy 341 ValCysTyrrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1171 GTGTGCTATGCTGGGCATGGACACTTGGAGAGGTGAGGCGAGTTACAGTGGCAAT 1230
 Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 Db 1231 ATCCAGGAGTTTGTGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC 1290
 Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 Db 1291 TTTGATGGGAGCCAGCCTCCACACTGCCCGCTCCAGCAGAGCAGCTCCAAGTGT 1350
 Qy 401 GluThrLeuGluGluIleThrGlyTyrrLeuTyrrIleSerAlaTrpProAspSerLeuPro 420
 Db 1351 GAGACTCTGGAAGAGATCAGAGTTACCTATACATCTCAGCATGGCCGAGCAGCCTGCCT 1410
 Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1411 GACCTCAGCTCTTCCAGAACCTGCAAGTAGTAATCGGGGAGCAATTTGCAACATGGCGCC 1470
 Qy 441 TyrrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1471 TACTCGTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCCTCACTGAGGAA 1530
 Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1531 CTGGGACAGTGGAGCCCTCATCCACCATACACCCACTCTGCTTCTGTGTCACACCGGTG 1590
 Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1591 CCCTGGGACAGCTCTTTCGGAACCCGACCAAGCTCTGTCTCCACACTGGCAACCGGCCA 1650
 Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 Db 1651 GAGGACAGTGTGTGGCGAGGCGCTGGCCCTGCCACAGCTGTGCGCCGAGGAGCACTGC 1710
 Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 Db 1711 TGGGTCCAGGGCCACCCAGTGTGTCACTGACAGCCAGTTCTTCCGGGCCAGGAGTGC 1770
 Qy 541 ValLeuGluCysArgValLeuGlnGlyLeuProArgGluTyrrValAsnAlaArgHisCys 560
 Db 1771 GTGGAGGAATGCCAGTACTGTCAGGGGCTCCCGCAGGAGTATGTGAATGCCAGCACTGT 1830
 Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1831 TTCCCGTGCACCCCTGAGTGTACAGCCCGCAGAAATGGCTCAGTACCTGTTTTGGACCGAG 1890
 Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrrLysAspProPheCysValAlaArgCys 600
 Db 1891 GCTGACAGTGTGGCTGTGCCACTATAGAGCCCTCCCTTCTGCGTGGGCCCGCTGC 1950
 Qy 601 ProSerGlyValLysProAspLeuSerTyrrMetProIleTrpLysPheProAspGluGlu 620


```

BASE COUNT      922 a 1382 c 1346 g      880 t
ORIGIN
Alignment Scores:
Pred. No.:      1.34e-186      Length:      4530
Score:          4892.00      Matches:      918
Percent Similarity: 73.15%      Conservative: 0
Best Local Similarity: 73.15%      Mismatches: 1
Query Match:      96.34%      Indels:      336
DB:              6      Gaps:      1

SEQ6 (1-919) x AR283481 (1-4530)

Qy      1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGluVala 20
Db      151 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCGCCCTTGGCCCCCGGAGCC 210
Qy      21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      211 GCGAGCACCAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 270
Qy      41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60
Db      271 ACCCACCCTGGACATGCTCCCGCACCTCTACAGGGCTGCCAGGTGGTGACAGGAAACCTG 330
Qy      61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db      331 GAATCACCCTACCTGCCACCAATGCCAGCTGTCTTCCTTCGAGGATATCCAGGAGGTG 390
Qy      81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      391 CAGGGCTACGTGCTCATCGCTCACAAACAGTAGGAGGCTCCACATGCAGAGCTGCGG 450
Qy      101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      451 ATTTGGCGAGCACCCAGCTCTTTGAGGACAACTATGCCCTTGGCCGTGCTAGACAAATGGA 510
Qy      121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      511 GACCCGCTGAACATACCAACCCCTGTACAGGGGCTTCCAGAGGAGGCTTCGGGAGCTG 570
Qy      141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db      571 CAGCTTGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG 630
Qy      161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      631 CTCTGCTACCAAGGACAGATTTTGTGAAGGACATCTTCCACAAGAACCAACAGCTGGCT 690
Qy      181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      691 CTCACATGATAGACACCAACCCGCTCTCGGGCTTGCACCCCTGTTCTCCGATGTGTAG 750
Qy      201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      751 GGCTCCCGCTGCTGGGAGAGAGTCTCAGGATTTTCAGAGCCTGACGCGCACCTGTCTGT 810
Qy      221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db      811 GCCGGTGGCTGTGCCCTGCAAGGGCCACTGCCCCACTGACTGCTGCCATGAGCAGTGT 870
Qy      241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      871 GCTCCCGGCTGCAGGGCCCCCAAGCACTCTGACTGCTGGCTGGCTTCCCTTCCACTTCAACAC 930
Qy      261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      931 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGCTACCTTACAAACACAGACACCTTTGAG 990
Qy      281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      991 TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGGCCAGCTGTGTGACTGTGCTGTCC 1050
```

```

Qy      301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      1051 TACAACCTACCTTTCTACGGAGCTGGGATCTTCACCCCTCGTCTGCTCCCTCGACACACCA 1110
Qy      321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      1111 GAGCTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTCAGCAAGCCCTGTGCCCGA 1170
Qy      341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1171 GTGTGCTATGCTTGGGCATGGACACTTGGAGAGGTGAGGGCAGTTACCAAGTGCCTAAT 1230
Qy      361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      1231 ATCCAGGAGTTTGTGCTGCAAGAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC 1290
Qy      381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      1291 TTTGATGGGACCCAGGCTCCAAACACTGCCCGCTCCAGCCAGCAGCAGCTCCAAGTGT 1350
Qy      401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db      1351 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGGCCGAGCAGCTGCCT 1410
Qy      421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db      1411 GACCTCAGCTCTTCCAGAACCTGCAGATTAATCGGGGACGAAATTCGCACAAATGGCGCC 1470
Qy      441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1471 TACTCGCTGACCCCTGAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCTCCTACTGAGGAA 1530
Qy      461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1531 CTGGGAGTGGACTGGCCCTCATCCACATAACACCCACCTGCTGCTGCTGCACACGCTG 1590
Qy      481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db      1591 CCCTGGGACCAAGCTCTTTGGGAACCCGACCAAGCTGCTGCTCCACTGCCACCGGCA 1650
Qy      501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db      1651 GAGGACGAGTGTGTGGCGAGGGCTGGCCCTGCCACAGCTGTGCGCCGAGGCGACATGC 1710
Qy      521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db      1711 TGGGGTCCAGGGCCCAACCTGCTCAACTGCAGCAGTTCTTTCGGGGCCAGGAGTGC 1770
Qy      541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db      1771 GTGGAGGAATGCCAGTACTGCAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1830
Qy      561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db      1831 TTGCGCTGCCACCTTGTGAGTGTGAGCCCAAGTGGCTGAGTGTGCTGTTGGACCGGAG 1890
Qy      581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db      1891 GCTGACCAAGTGTGGGCTGTGCCCTATTAAGGACCTTCCCTTCTGCGTGGCCGCTGC 1950
Qy      601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db      1951 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
Qy      621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db      2011 GGGCGATGCCAGCTTGCCTCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 2070
Qy      641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
Db      2071 GGCTGCCCCCGGAGCAGAGAGCCAGCCCTCTGACGTCCATCGCTCTGCGGTGGTGGC 2130
```


QY 653 ----- 653
 Db 2131 ATTCTGCTGCTGCTGCTTTGGGGTGGTCTTTTGGGATCCTCATCAAGCAGCGGACAG 2190
 QY 653 ----- 653
 Db 2191 AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGAGCTGTGGAGCGCTG 2250
 QY 653 ----- 653
 Db 2251 ACACCTAGCGGAGCGATGCCAACCCAGCGCAGATGCCGATCCTGNAAGAGAGGAGCTG 2310
 QY 653 ----- 653
 Db 2311 AGGAAGTCAAGGTGCTTGATCTGGCATCTGGCAGTCTTTGGCACAGTCTACAAGGGCATCTGGATC 2370
 QY 653 ----- 653
 Db 2371 CCTGATGGGAGATGTGAAATTCAGTGGCCATCAAAGTGTTCAGGGAAACACATCC 2430
 QY 653 ----- 653
 Db 2431 CCCAAGCCACAAAGAAATCTTAGACGAAGCATACGTCATGCTGGTGGTGGCTCCCA 2490
 QY 653 ----- 653
 Db 2491 TATGTCTCCGCTTCTGGGCATCTGCCTGACATCCACGGTGACAGCTT 2550
 QY 653 ----- 653
 Db 2551 ATGCCATATGGTGCCTTTAGACCATGTCCCGGAAACCGGAGCGCTGGGCTCCAG 2610
 QY 653 ----- 653
 Db 2611 GACCTCCTGAAGTGTATGCAGATTGCCAAGGGATGAGTACCTGAGGATGTGGG 2670
 QY 653 ----- 653
 Db 2671 CTCGTACACAGGACTTGGCCGTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA 2730
 QY 653 ----- 653
 Db 2731 ATTACAGACTTCGGCTGCTCGGCTGTGACATTGACGACAGATACCATGCAGAT 2790
 QY 653 ----- 653
 Db 2791 GGGGGCAAGTGCCCATCAAGTGGATGGCGTGGAGTCCATTCTCCGCGCGGTTCACC 2850
 QY 653 ----- 653
 Db 2851 CACCAGAGTATGTGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2910
 QY 653 ----- 653
 Db 2911 AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2970
 QY 653 ----- 653
 Db 2971 CTGCCCCAGCCCCCATCTGCACCATTTGATGCTTACATGATCATGGTCAATTTGGATG 3030
 QY 653 ----- 653
 Db 3031 ATTGACTCTGAATGTGGCCAAAGATTCGCGGAGTTGGTGTCTGAATTCCTCCCGATGGCC 3090
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 Db 3091 AGGAGCCCCCAGCGCTTTGTGTATCCAGATGAGGACTTGGGCCACCCAGTCCCTTG 3150
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 Db 3151 GACAGACCTTCTACCGCTCATCTGCTGGAGGACGATGACATGGGGACCTGTGTGATGCT 3210
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAlaProGlyAlaGly 704

Db 3211 GAGGAGTATCTGTTACCCAGCAGGGCTTCTTCTGTCCAGACCTGCCCGGGCGCTGGG 3270
 QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
 Db 3271 GGCATGGTCCACACAGGACCGCAGCTCATCTACCAGAGTGGGGTGGGACCTTGACA 3330
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
 Db 3331 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCGCAGGTCTCCACTGGCACCTCCGAAGG 3390
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
 Db 3391 GCTGGCTCGATGATTTGATGTCACCTGGGAATGGGGCAGCCAAAGGGGTGCAAGC 3450
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
 Db 3451 CTCCCACACATGACCCCGCAGCCTCTACAGCGGTACAGTGGAGACCCACATACCCCTG 3510
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
 Db 3511 CCCTCTGAGACTGATGGCTACGTTGCCCGCTGACCTGCAGCCCGCAGCTGAATATGTG 3570
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
 Db 3571 AACCAAGCCAGATGTTGGGCCCGCAGCCCTTGGCCCGCAGAGGGGCTTGCCTGCTGCC 3630
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
 Db 3631 CCACCTGCTGGTGGCCTCTGGAAGGGCCAGACTCTCTCCCGAGGAGAAATGGGGTTC 3690
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
 Db 3691 GTCAAGAGCTTTTGCCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3750
 QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
 Db 3751 GGAGGAGCTGCCCCCTCAGCCCGCAGCTCTCTCTCTCTCAGCCCGCAGCTTCGACAACTC 3810
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
 Db 3811 TATTACTGGGACAGGACCCAGAGCGGGGGTCTCCACCCAGCAGCTTCAAAAGGACA 3870
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db 3871 CCTACGGCAGAACCCAGAGTACCTGGGTCTGGAGCTGCCAGTG 3915
 RESULT 14
 AX282577
 LOCUS AX282577 4530 bp DNA linear PAT 02-NOV-2001
 DEFINITION Sequence 1 from Patent WO0168677.
 ACCESSION AX282577
 VERSION AX282577.1 GI:16609668
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Nicolette,C.A.
 TITLE Therapeutic compounds
 JOURNAL Patent: WO 0168677-A 1 20-SEP-2001;
 GENZYME CORPORATION (US)
 FEATURES
 source Location/Qualifiers
 1..4530
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 151..3918
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD10597.1"
 /db_xref="GI:16609668"
 CDS

/translation="MELAALCRWGLLIALLPPGAASQTQVCTGTDKMLRLRLPASPEETHLD
MLRHLYGQWOGNLELTYLPNTNASIFLDIQEVOGYVYLIAHNOVROVQLORLIV
RCTQLDEGDNVALVDNGDPLNPTTPTWASPGGLRELQRLSLEILKGVLVQRNPQ
LCVQDTILKWDIFHKNNQALATLIDTNRSRACHPCSPMKAGSRGWESSEDCSLKRT
VCAGGACRCKGPIPTDCHCQACAGCTGPKHSDCLACLFNHSGIOELHCPALVTYNT
DTPESNPEGRYTFGASCCTACGPNYLTSDVGSCTLVCPQHNOEVTAEQTCRCEK
SKPCARVCYGLGHEHLEVRVTSANTIOEPAGCKIFGSLAFLPESFDGDSQNTAPL
QPOLQVFEETLEETGYLISAWPDSLPDLVSFQNLQVIRGRILHNCAYSLTLQGLGI
SWLGRCLDHVSGSLALIHNTHLICFVHTVPMDLPFNHQAALLHNRPEDECEGEG
LACHQLCARHGCGPQTQVNCVSQFLRGQECVEECRVLQGLPREVYNARHCLCPHE
CQONGSVTCFGEADQCVACAHYKOPFCVAKCPGVKPDLSPTLWKPFPDEEGACQ
PCPINTHSCVDLDDGCPAEQASPLSTISVAWGLLVGLVGVFGILIKRRQKI
RYKTMRLLEQTELVEPLTPSGAMPNOAMRILKETELRVKVLGSGAFGTVYKGLWI
PDGENKPIVAIKRENTSPKANKETLDEAYVMAGVSGSPVSRLLGLCLTSTVQLVT
OLMPYGLDHLVRENKRLGSDLLNMCIOIAKMSYLEDVRLVRLDARNLVQSP
NHVKITDFGLARLLDDIETEHADGKVPYIKWMALESILLRRTHOSDWSYGVYWE
LMTFGAKPDGTFAREIPDLLEGERPQPICTIDYIMIMVKWMDISCRPRFREL
VSEFSRNARDQRFVYIQLNEDLGPASPLDSTFYRSLEDDMDGDLDAEYLYPQQGF
FCPDPAAGAGMYHHRSSSTRSGGDLTLGLEPSEEEAPRSP LAPSEGAGSDVFDG
DLGMGAAGLQSLPTHDPSLPQRYSDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVR
POPSPREGPLPAARAGATLERAKTLLSPCKNGVWVDVFAFGAVENPEYLTPOGGAA
POPHPPAFSPADNLYWDDPPERGAPSTFKGTPTAENPEYLGLDYVPV"

BASE COUNT 922 a 1382 c 1346 g 880 t
ORIGIN

Alignment Scores:

Pred. No.: 1,34e-186 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservations: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 6 Gaps: 1

SEQ6 (1-919) x AX282577 (1-4530)

Qy	1	MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla	20
Db	151	ATGGAGCTGGCGGCTTGTCCGCTGGGGCTCCTCTCTCCCTCTTGTCCGCCCGGAGCC	210
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	211	GCAGACCCCAAGTGTGCACCGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG	270
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu	60
Db	271	ACCCACCTGGACATGCTCCGCCACCTCTACACAGGCTGCCAGTGGTGCAGGGAACCTG	330
Qy	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	331	GAACCTACCTACCTGCCGCCACCAATGCCAGCCTGTCTCTTCTTCCAGGATATCCAGGAGTG	390
Qy	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	391	CAGGCTACCTGTCTCATCGCTACAAACCAAGTAGGAGGCTCCACCTGCAGAGGCTCGG	450
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	451	ATTGTGGAGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACATGGA	510
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyLeuArgGluLeu	140
Db	511	GACCCGCTGAACAATACACCCCTGTCCAGGGGCCCTCCACAGAGGCTCGCGGAGCTG	570
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
Db	571	CAGCTCGAAGCTTCACAGAGATCTTGAAAGGAGGCTGTGATCCAGCGGAACCCCCAG	630
Qy	161	LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	631	CTCTGCTACCCAGGACGATTTGTGGAAGACATCTTCACAGAACAACACAGCTGGCT	690
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	1771	GTGGAGGAATGCCAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1830
Db	691	CTCACACTGATAGACACCAACCGCTCTCGGGCCTGCACCCCTGTTCTCCGATGTGTAAG	750
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	751	GGTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTTGAGAGCCTGCAGCGCCTGCTGTGT	810
Qy	221	AlaGlyGlyCysAlaAraGlyCysGlyProLeuProThrAspCysCysHisGlnGlnCys	240
Db	811	GCCGCTGGCTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGACGACTGT	870
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	871	GCTGCCGCTGCACGGGCCCAAGACACTCTGACTGCTGCCCTGCCCTGCCCTCAACAC	930
Qy	261	SerClyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	931	AGTGGCATCTGTGAGCTGCACTGCCACGCCCTGCTACCTACCAACACAGACACACTT	990
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	991	TCCATGCCCAATCCCGAGGGCCGATATACATTGCGGCCAGCTGTGTGACTGCTGCTCC	1050
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1051	TACAACTACCTTTCTACGGACGTGGGATCCTGCACCTCTGCTGCCCTGCACACCA	1110
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1111	GAGGTGACACAGAGGATGAACACACGCGGTGTGAGAAGTGCACAGACCCCTGTGCCGA	1170
Qy	341	ValCystTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1171	GTGTCTATGCTTGGCATGGGACCTTCCGAGAGGTGAGGCGAGTACCAGTGCCTAAT	1230
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1231	ATCCAGGAGTGTGCTGGCTGCAAGAGATCTTTGGGAGCCTGGCATTTCTGCCGAGAGC	1290
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnGlnValPhe	400
Db	1291	TTTGTATGGGACCCAGCCTCCACACTGCCCGCTCCAGCCAGCAGAGCTCCAAGTGT	1350
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1351	GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGACTGCCCGGACAGCCTGCCT	1410
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1411	GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATTCGCAATGGGCGC	1470
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1471	TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGGCTGCTACTGAGGAA	1530
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1531	CTGGCAGCTGAGTGGCCCTCATCCACCATACACCCACCTCTGCTTCGTCCGACACCGTG	1590
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1591	CCCTGGACCAAGCTCTTTCGGAACCCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1650
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1651	GAGACAGAGTGTGGGCGAGGGCTGCGCTGCCCTGCCAGCTGTGCGGCCGAGGCACTGC	1710
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1711	TGGGTCACGGGCCACCCAGTGTGCTCACTGCAGCAGCTTCCTTCGCGGCGCAGGAGTGC	1770
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAlaAlaArgHisCys	560
Db	1771	GTGGAGGAATGCCAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1830

```
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1831 TTGCCGTGCCACCTTGAGTGTACAGCCCAAGATGCTCAGTACCTGTTTGGACCGGAG 1890
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1891 GCTGACAGTGTGTGGCTGTGCCACTATAAGGACCTCCCTTCTGCGTGGCCGCTGC 1950
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu 620
Db 1951 CCAGGGGTGTGAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 2011 GCGCATGCCAGCTTGCCTCCATCACTGCACCCACTCTCTGTGTGACCTGTATGACAG 2070
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer- 653
Db 2071 GGCTGCCCGCGAGCAGAGAGCCAGCCCTCTGAGTCCATGCTCTCTGCGGTGTTGCG 2130
QY 653 ----- 653
Db 2131 ATTCTGCTGGTGTGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGAGCGGAGCAG 2190
QY 653 ----- 653
Db 2191 AAGATCCGGAAGTACACGATGCGAGAGTGTCTGCAGAAACGGAGCTGTGTGGAGCGCTG 2250
QY 653 ----- 653
Db 2251 ACACCTAGCGGAGCGATGCCAACCAAGCGCGAGATCGGATCCTGAAAGAGACGGAGCTG 2310
QY 653 ----- 653
Db 2311 AGGAAGTGAAGTGTGGATCTGGCGTCTTTGGCACAGTCTACAAGGGCATCTGGATC 2370
QY 653 ----- 653
Db 2371 COTGATGGGAGAAATGTAAATTCAGTGCCATCAAAGTGTGAGGAAACACATCC 2430
QY 653 ----- 653
Db 2431 CCCAAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGGCTGTGTGGCTCCCCA 2490
QY 653 ----- 653
Db 2491 TATGTCCTCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGTGTGACACAGCTT 2550
QY 653 ----- 653
Db 2551 ATGCCCTATGCTGCTCTTAGACCATGTCCGGGAAACCGGAGCGCTGGCTGCCCA 2610
QY 653 ----- 653
Db 2611 GACCTGCTGAAGTGTGTATGCAGATTGCCAAGGGGATGAGTACCTGGAGGATGTGCGG 2670
QY 653 ----- 653
Db 2671 CTCGTACAGGGACTTGGCCSCTCGGAACGTGTGGTCAAGAGTCCCACACCATGTCAA 2730
QY 653 ----- 653
Db 2731 ATTACAGACTTCGGGTGCTCGGCTGTGGATTTGACATTGACGACAGAGTACCATGCAGAT 2790
QY 653 ----- 653
Db 2791 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGCTTCCAC 2850
QY 653 ----- 653
Db 2851 CACCAGATGATGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2910
```

```
QY 653 ----- 653
Db 2911 AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGCTGAAAGGGGAGCGG 2970
QY 653 ----- 653
Db 2971 CTGCCCCAGCCCCCATCTGSCACCATTTGATGTCTACATGATCATGTCACAAATGTTGGATG 3030
QY 653 ----- 653
Db 3031 ATTGACTCTGAATGTGGCCCAAGATTCGGGAGTTGGTCTCTGAATTTCTCCCGCATGGCC 3090
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 3091 AGGAGCCCCACGCGCTTTGTGTCTATCCAGATAGGACTTGGGCCACCCAGTCCCTTG 3150
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACCTTCTACCGCTCACTGTGTGGAGACGATGACATGGGACCTGGTGGATGCT 3210
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3211 GAGGAGTATCTGTATCCCGCAGCGGCTTCTTCTGTCCAGACCTGCCCGGGCGCTGGG 3270
QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
Db 3271 GGCATGGTCCACACAGGCACCGCAGCTCATCTACCAGAGTGGCGTGGGACCTGACA 3330
QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3331 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCGCCAGGTCTCCACTGGCACTCCCGAAGG 3390
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3391 GCTGGCTCCAGTATTTGATGGTACCTGGGAATGGGGCGACCCAGGGGTGCCAAGC 3450
QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3451 CTCCCACACATGACCCCGACCCCTCTACAGCGGTACAGTGAGGACCCACAGTACCCCTG 3510
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3511 CCCTCTGAGACTGATGGCTACGTTGCCCTTGCCTGACCTGCAGCCCGCAGCTGAATATG 3570
QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3571 AACAGCCAGATGTTTCGGCCCGACCCCTTCCGCCCGAGAGGGCCCTCTGCTGCTGCC 3630
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3631 CGACCTGCTGTGCCACTCTGGAAAGGGCCAAAGACTCTCTCCCAGGGAAGAATGGGTC 3690
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3691 GTCAAAGAGTGTGCTTGGGGTGGCGGAGAACCCGAGTACTTGACACCCAG 3750
QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3751 GGAGGAGTGGCCCTCAGCCCGACCCCTCTCTGCTTCCAGCCCGACCTTCGACACCTC 3810
QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3811 TATTACTGGGACGAGCCACGAGCGGGGGTCCACCCAGCAGCTTCAAAGGAGCA 3870
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3871 CCTACGGCAGAACCCAGAGTACCTGGGTCTGGAGCTGGCCAGTG 3915
RESULT 15
AX587649
LOCUS
DEFINITION Sequence 119 from Patent WO0246467.
ACCESSION AX587649
```

4530 bp DNA linear PAT 10-JAN-2003

VERSION	AX587649.1	GI:28212358
KEYWORDS	synthetic construct	
SOURCE	synthetic construct	
ORGANISM	artificial sequences.	
REFERENCE	1	
AUTHORS	Bertucci,F., Houlgatte,R., Birnbaum,D., Nguyen,C., Viens,P. and Fert,V.	
TITLE	Gene expression profiling of primary breast carcinomas using arrays of candidate genes	
JOURNAL	Patent: WO 0246467-A 119 13-JUN-2002;	
FEATURES	Ipsogen (FR)	
source	Location/Qualifiers	
	1..4530	
	/organism="synthetic construct"	
	/db xref="taxon:32630"	
	/note="primer"	
misc_feature	1..4530	
	/note="v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2 (neuro/glioblastoma derived oncogene homolog) (ERBB2) gene."	
BASE COUNT	922 a	1382 c 1346 g 880 t
ORIGIN		
Alignment Scores:		
Pred. No.:	1.34e-186	Length: 4530
Score:	4892.00	Matches: 918
Percent Similarity:	73.15%	Conservative: 0
Best Local Similarity:	73.15%	Mismatches: 1
Query Match:	96.34%	Indels: 336
DB:	6	Gaps: 1
SEQ6 (1-919) x AX587649 (1-4530)		
Qy	1	MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuProGlyAla 20
Db	151	ATGGAGTGGCGCTTGTCCGCTGGGGCTCTCTCTCGCCCTCTGCCCGGAGCC 210
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db	211	GCAGACCCAGTGTGCACGGGCACAGACATGAGCTGGGGCTCCCTGCCAGTCCCGAG 270
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyfGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db	271	ACCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGTGTGACGGAAACCTG 330
Qy	61	GluLeuThrTyfLeuProThrAsnAlaSerPheLeuGlnAspIleGlnGluVal 80
Db	331	GAATCTACCTACCTGGCCACCAATGCCAGCTGTCTCTCTGCAGGATATCCAGGAGGTG 390
Qy	81	GlnGlyTyfValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db	391	CAGGGTACGTGTCTCATCGTCAACCAAGTACGAGGAGGTCCCATCTGCAGAGGTGCGG 450
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyfAlaLeuAlaValLeuAspAsnGly 120
Db	451	ATTGTGGAGGCACCCAGCTTCTTGAGGACAACTATGCTCCCTGGCGCTGTAGACAATGGA 510
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db	511	GACCCGTGAACAATACACCCCTGTCTACAGGGGCTCCCGCAGGAGGCTGTCCGGAGCTG 570
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db	571	CAGCTTCGAAGCTTCACAGAGATCTTGAAGAGGGGTCTGTATTCACGGCGAACCCTCCAG 630
Qy	161	LeuCystTyfGlnAspThrIleLeuTrpTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db	631	CTCTGCTACAGGACACGATTTTGTGGAAGGACATCTTCCACAAAGAACACACAGCTGGCT 690
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db	691	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCTCGATGTGTGAAG 750

QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1831	TTGCCGTGCCACCTGAGTGTACAGCCCAAGATGGCTCAGTCACTGTTTGGACCGGAG	1890
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1891	GCTGACCACTGTGGCCCTGTGCCCACTATAAGGACCCCTCCCTTCTGCGTGGCCGCTGC	1950
QY	601	ProSerGlyValLysProAspLysSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1951	CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2010
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Db	2011	GGCGCATGCCAGCCTTGCCCACTCAACTGCACCCACTCTCTGTGTGACCTGGATGACAAG	2070
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2071	GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCGTCTCTCGGTGGTTGGC	2130
QY	653	-----	653
Db	2131	ATTCTCCTGCTGCTGCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAG	2190
QY	653	-----	653
Db	2191	AGATCCGGAAGTACACGATGGGAGACTGCTGCAGAAACGGAGCTGGTGGAGCGCTG	2250
QY	653	-----	653
Db	2251	ACACCTAGCGGAGCGATGCCCAACACAGGCGCAGATCGGATCGTGAAGAGACGGAGCTG	2310
QY	653	-----	653
Db	2311	AGGAAGTGAAGTGTGATCTGGCGTCTTTGGCACAGCTCTACAAGGGCATCTGGATC	2370
QY	653	-----	653
Db	2371	CCTGATGGGAGAAATGTGAATTTCCAGTGGCCATCAAAAGTGTGAGGAAACACATCC	2430
QY	653	-----	653
Db	2431	CCCAAGCCCAACAAAGAAATCTTAGACGAAGCATACGTATGGCTGGTGGGCTCCCCA	2490
QY	653	-----	653
Db	2491	TATGTCTCCGCGCTTCTGGGCATCTGCCTGACATCCACGTCGACGCTGGTGACACAGCTT	2550
QY	653	-----	653
Db	2551	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCAC	2610
QY	653	-----	653
Db	2611	GACCTGCTGACTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2670
QY	653	-----	653
Db	2671	CTCGTACACAGGACTTGGCCGCTCGGAAGCTGCTGTCGAAGATCCCAACCATGTCAAA	2730
QY	653	-----	653
Db	2731	ATTACAGACTTCGGGTGGCTCGGCTGCTGGACATTGACGAGACAGATACCATGCAGAT	2790
QY	653	-----	653
Db	2791	GGGGCAAGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGCGGTTACCC	2850
QY	653	-----	653
Db	2851	CACCAGATGATGTGGAGTTATGGTGTACTGTGTGGAGCTGATGACTTTTGGGCC	2910
QY	653	-----	653
Db	2911	AAACCTTACGATGGATCCACCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG	2970
QY	653	-----	653
Db	2971	CTGCCCCAGCCCCCATCTGCACCATTTGATGTTCTACATGATGTTCAATGTTGATG	3030
QY	653	-----	653
Db	3031	ATTGACTCTGAATGTCCGCCAAGATTCGGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC	3090
QY	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	3091	AGGGACCCCAAGCGCTTTGTGTCATCCAGAATAGAGACTTGGGCCACGCCAGTCCCTTG	3150
QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3151	GACAGCACCTTCTACCGCTCAGCTGCTGGAGGACATGACATGGGGACCTGGTGGATGCT	3210
QY	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3211	GAGGACTATCTGTTACCCAGCAGGCGCTTCTTCTGTCTCAGACCTGCCCGGCGCTGG	3270
QY	705	GlyMetValHisHisArgSerSerSerThrArgSerGlyGlyAspLeuThr	724
Db	3271	GGCATGGTCCACACAGGACCGCAGCTCATCTACCAGGAGTGGCGTGGGACCTGACA	3330
QY	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	3331	CTAGGCGCTGGAGCCCTCTCAAGAGAGGCCCCAGGTCTCCACTGGCACCTCCGAAAGG	3390
QY	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
Db	3391	GCTGGCTCCGATGATTTGATGTGACCTGGGAATGGGGACGCCAAGGGCTGCAAGC	3450
QY	765	LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu	784
Db	3451	CTCCACACATGACCCCGCCCTCTACAGCGGTACAGTGAGGACCGACAGTACCCCTG	3510
QY	785	ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal	804
Db	3511	CCCTCTGACACTGATGGCTAGCTGGTGGCCCTGACCTGCAGCCCGCTGAATATGTG	3570
QY	805	AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla	824
Db	3571	AACCAGCCAGATCTTCGGCCCGCCAGCCCTTCGCCCGAGAGGGCCCTCTGCTGTGTC	3630
QY	825	ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal	844
Db	3631	CGACCTGCTGGTCCCACTCTGGAAGGGGCCAAGACTCTCTCCCGAGGAAGATGGGTC	3690
QY	845	ValLysAspValPheAlaPheGlyGlyAlaValAlaGluAsnProGluTyrLeuThrProGln	864
Db	3691	GTCAAGACGCTTTTGGCTTGGGGTGGCGTGGAGAACCCCGACTTGTGACACCCAG	3750
QY	865	GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu	884
Db	3751	GGAGGAGCTGCCCTCAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3810
QY	885	TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr	904
Db	3811	TATTACTGGGACCAAGGACCCAGAGCGGGGGGCTCCACCCAGCACTTCTCAAGGACA	3870
QY	905	ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal	919
Db	3871	CCTACGCGAGAAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG	3915

Search completed: October 16, 2003, 11:05:33
Job time : 13378.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 461.761 Seconds
(without alignments)
3105.423 Million cell updates/sec

Title: SEQ5
Perfect score: 326
Sequence: 1 QNEDLGASPLDSTFYRSLL.....GFFCPDPAGGGMVHRRH 59

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cyn2_1/USPPO_Spool_P/HOLLERAN480/runat_15102003_131913_20548/app_query.fasta_1.4685
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGN_1_12645=runat_15102003_131913_20548 -NCPU=3
-NO_WMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssi:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	326	100.0	378	14	CB266492	CB266492 1005398 H
2	326	100.0	423	12	BM703962	BM703962 UI-E-CK1-
3	326	100.0	464	2	HSM067221	Bx477782 Homo sapi
4	326	100.0	531	12	BM830170	BM830170 K-EST0103
5	326	100.0	536	14	CB129403	CB129403 K-EST0179
6	326	100.0	547	12	BM787824	BM787824 K-EST0066
7	326	100.0	574	12	BM829991	BM829991 K-EST0103
c	326	100.0	691	14	CB853376	CB853376 UI-CF-FNO
	326	100.0	849	14	CA489799	CA489799 AGENCOURT
9	326	100.0	852	14	CD516283	CD516283 AGENCOURT
10	326	100.0	885	14	CA455074	CA455074 AGENCOURT
11	326	100.0	893	14	CA455141	CA455141 AGENCOURT
12	326	100.0	894	14	CA454131	CA454131 AGENCOURT
13	326	100.0	899	14	CD558386	CD558386 AGENCOURT
14	326	100.0	916	13	BU845369	BU845369 AGENCOURT
15	326	100.0	919	13	BU845391	BU845391 AGENCOURT
16	326	100.0	932	14	CA487981	CA487981 AGENCOURT
17	326	100.0	944	14	CD558559	CD558559 AGENCOURT
18	326	100.0	1004	14	CD515356	CD515356 AGENCOURT
19	326	100.0	4715	11	AF318349	AF318349 Homo sapi
20	322	98.8	685	9	AU123871	AU123871 AU123871
21	313	96.0	924	14	CA489100	CA489100 AGENCOURT
22	313	96.0	984	14	CA489084	CA489084 AGENCOURT
23	309	94.8	866	14	CA488343	CA488343 AGENCOURT
24	302	92.6	348	9	A1909847	A1909847 QV-BT225-
25	300	92.0	444	12	BM789286	BM789286 K-EST0068
26	300	92.0	518	10	BF039952	BF039952 BP250015A
27	300	92.0	537	12	BM793744	BM793744 K-EST0074
28	295	90.5	505	12	BM854421	BM854421 K-EST0136
29	292	89.6	577	13	BQ831943	BQ831943 LL6in1122
30	283	86.8	294	10	BF757857	BF757857 CM4-CT057
31	283	86.8	674	10	BG675003	BG675003 602621225
32	279	85.6	527	14	CB719751	CB719751 AMGNNUC:N
33	277	85.0	496	12	BM795191	BM795191 K-EST0076
34	276	84.7	422	10	BF554233	BF554233 UI-R-CO-h
35	276	84.7	904	13	BU931894	BU931894 AGENCOURT
36	275	84.4	630	2	HSM073298	Bx483130 Homo sapi
37	269	82.5	568	13	BQ831943	BQ831943 LL6in1122
38	269	82.5	577	13	BQ829915	BQ829915 LL6in2239
39	269	82.5	592	10	BE382065	BE382065 601272225
40	269	82.5	629	10	BE374596	BE374596 601223384
41	269	82.5	715	13	BU052728	BU052728 UI-M-EXO-
42	269	82.5	726	12	BI648082	BI648082 603279290
43	269	82.5	770	12	BI645946	BI645946 603274921
44	269	82.5	773	12	BI157032	BI157032 602921306
45	269	82.5	785	12	BI688650	BI688650 603311262

ALIGNMENTS

RESULT 1	CB266492	1005398 Human Fat Cell 5'-Stretch Plus cdna Library Homo sapiens	378 bp	mrna	linear	EST 20-FEB-2003
CB266492	LOCUS	CDNA 5', mRNA sequence.				
DEFINITION	CB266492	EST.				
ACCESSION	CB266492.1	GI:28441078				
VERSION	EST.					
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 378)					

AUTHORS **TITLE** **JOURNAL** **COMMENT**

Yang, R.-Z., Shuldiner, A. and Gong, D.-W.
EST analysis of human adipose gene expression
Unpublished
Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition
University of Maryland
660 Redwood St, HH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgong@medicine.umaryland.edu
PCR primers
FORWARD: CTCGGGAAGCGCCCATGTTGTGGT
BACKWARD: AATAGACTACTATAGCGCAATTGG
Seq primer: GTTGGTACCGGGAATC.

FEATURES

source

1..378
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/note="Vector: lambda triplex"

BASE COUNT 71 a 110 c 117 g 80 t

ORIGIN

Alignment Scores:
Pred. No.: 1.72e-29 Length: 378
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

SEQ5 (1-59) x CB266492 (1-378)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 133 CAGAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTCTACCGCTCAGTGTG 192
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
|||||
Db 193 GAGGACGATGACATGGGGACCTGTGGATGCTGAGGAGTATCTGTGTACCCAGCAGGCG 252
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
Db 253 TTCTTCTGTCCAGACCTGCGCGCGCTGGGGGCGATGTCACACAGGACCCGC 309

RESULT 2

BM703962

LOCUS

DEFINITION BM703962 423 bp mRNA linear EST 28-FEB-2002
UI-E-CK1-afk-o-11-0-UI-ri UI-E-CK1 Homo sapiens cDNA clone
UI-E-CK1-afk-o-11-0-UI 5', mRNA sequence.

ACCESSION BM703962

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 423)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEHRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

FEATURES

source

1..423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="UI-E-CK1-afk-o-11-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-CK1 is a normalized cDNA library containing the
following tissue(s): Retina Foveal and Macular. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GTCC. This library was created for the program,
Gene Discovery in the Visual System, supported by National
Gene Institute (NEI)." 85 t

BASE COUNT

78 a 129 c 131 g

ORIGIN

Alignment Scores:

Pred. No.: 1.95e-29 Length: 423

Score: 326.00 Matches: 59

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

SEQ5 (1-59) x BM703962 (1-423)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20

Db 156 CAGAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTCTACCGCTCAGTGTG 215

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40

Db 216 GAGGACGATGACATGGGGACCTGTGGATGCTGAGGAGTATCTGTGTACCCAGCAGGCG 275

QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59

Db 276 TTCTTCTGTCCAGACCTGCGCGCGCTGGGGGCGATGTCACACAGGACCCGC 332

RESULT 3

HSN067221

ID HSN067221 standard; RNA; EST; 464 BP.

XX BX477782;

AC BX477782.1

XX SV

XX DX

DT 09-MAY-2003 (Rel. 75, Created)

DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

XX Homo sapiens mRNA; EST DKFZp686M12198_r1 (from clone DKFZp686M12198)

DE Homo sapiens mRNA; EST DKFZp686M12198_r1 (from clone DKFZp686M12198)

XX Homo sapiens mRNA; EST DKFZp686M12198_r1 (from clone DKFZp686M12198)

KW EST; expressed sequence tag.

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers


```

XX Homo sapiens (human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX [1]
RN 1-464
RP Bloecker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.
RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
XX This is the 5' sequence of the clone insert
CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC Sequenced by GBF (National Research Centre for Biotechnology
CC Ltd., Braunschweig/Germany) within the cDNA sequencing
CC consortium of the German Genome Project.
CC No sl sequence available.
CC This clone (DKFZp686M12198) is available at the RZPD in Berlin.
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX FH Location/Qualifiers
FH 1. 464
FT /db_xref="taxon:9606"
FT /mol_type="mRNA"
FT /organism="Homo sapiens"
FT /clone="DKFZp686M12198"
FT /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
FT DH10B; sites SfiIA + SfiIB"
FT /dev_stage="adult"
FT /tissue_type="cDNA-collection"
XX Sequence 464 BP; 87 A; 127 C; 149 G; 101 T; 0 other;

Alignment Scores:
Pred. No.: 2,15e-29 Length: 464
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

SEQ5 (1-59) x HSM067221 (1-464)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
Db 257 CAGATGAGGACTTGGGCCACCCAGTCCCTGGGACAGCACCCTTCTACCGCTCAGTGTG 316
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyLeuValProGlnGlnGly 40
Db 317 GAGGACGATGACATGGGGACCTGGTGGTGTGAGGAGTATCTGGTACCCAGCAGGCG 376
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 377 TTCCTTGTCCAGACCTTGGCCCGGCGCTGGGGCGATGGTCCACACGACCGCC 433

RESULT 4
BM830170
LOCUS BM830170 531 bp mRNA linear EST 06-MAR-2002
DEFINITION K-EST0103592 S21SNU520s1 Homo sapiens cDNA clone S21SNU520s1-19-D03
5', mRNA sequence.
ACCESSION BM830170
VERSION BM830170.1 GI:19186579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Hominidae; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 531)

```

```

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 19 row: D column: 03
High quality sequence stop: 531.
Location/Qualifiers
1. 531
Source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNU520s1-19-D03"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/lab_host="SNU-520"
/clone_lib="S21SNU520s1"
/Note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 ~ 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(dT)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the subtracted cDNA
libraries were constructed by transformation of the
remaining DNA into competent cells E. coli Top10F' with
electroporation method."
BASE COUNT 102 a 149 c 166 g 114 t
ORIGIN

Alignment Scores:
Pred. No.: 2,5e-29 Length: 531
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ5 (1-59) x BM830170 (1-531)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
Db 343 CAGATGAGGACTTGGGCCACCCAGTCCCTGGGACAGCACCCTTCTACCGCTCAGTGTG 402
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyLeuValProGlnGlnGly 40

```

```

|||||
403 GAGGACCATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGC 462
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
463 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGGCGATGCTCCACACAGGCCCGC 519

RESULT 5
LOCUS CB129403 536 bp mRNA linear EST 29-JAN-2003
DEFINITION K-EST0179037 C1SNUI7 Homo sapiens cDNA clone C1SNUI7-5-C08 5', mRNA
sequence.
ACCESSION CB129403
VERSION CB129403.1 GI:28093000
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 536)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 5 row: C column: 08
High quality sequence stop: 536.
Location/Qualifiers
1. 536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="C1SNUI7-5-C08"
/sex="F"
/tissue_type="Uterine"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="C1SNUI7"
/notes="Organ: Cervix; Vector: pCNS-D2; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dT-selected mRNA by
priming with dr-tailed vector. The dr-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
BASE COUNT 105 a 150 c 167 g 114 t
ORIGIN

Alignment Scores:
Pred. No.: 2,52e-29 Length: 536
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

```

```

SEQ5 (1-59) x CB129403 (1-536)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 348 CAGAAATGAGACTTGGCCCGCAGTCCCTTGACAGCACCCTTACCCCTCACTGCTG 407
|||||
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 408 GAGGACGATCACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGC 467
|||||
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
Db 468 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGGCGATGCTCCACACAGGCCCGC 524

RESULT 6
LOCUS BM787824 547 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0066898 S1LSNU1 Homo sapiens cDNA clone S1LSNU1-23-G01 5',
mRNA sequence.
ACCESSION BM787824
VERSION BM787824.1 GI:19136056
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 23 row: G column: 01
High quality sequence stop: 547.
Location/Qualifiers
1. 547
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1LSNU1-23-G01"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/lab_host="Top10F"
/clone_lib="S1LSNU1"
/notes="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested pME18S-FL3 vector. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
BASE COUNT 102 a 187 c 163 g 95 t
ORIGIN

Alignment Scores:
Pred. No.: 2,58e-29 Length: 547

```

Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

SEQ5 (1-59) x BM87824 (1-547)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 |||||
 Db 48 CAGAAATGAGGACTTGGGCCCGCCAGCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 107
 |||||
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
 |||||
 Db 108 GAGGACGATGACATGGGGGACCTGGTGATCTGAGGAGTATCTGGTACCCACGACGGC 167
 |||||
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
 |||||
 Db 168 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGTCTCCACACGACGACCGC 224
 |||||

RESULT 7

BM829991
 LOCUS K-EST0103390 S21SNU520s1 Homo sapiens cDNA clone S21SNU520s1-17-C05
 DEFINITION 5', mRNA sequence.

ACCESSION BM829991.1 GI:19186400

VERSION BM829991

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 574)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished

COMMENT Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@email.kribb.re.kr

Plate: 17 row: C column: 05

High quality sequence stop: 574.

Location/Qualifiers

1..574

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/cclone="S21SNU520s1-17-C05"

/sex="F"

/tissue_type="Stomach"

/cell_type="Floating aggregates"

/cell_line="SNU-520"

/lab_host="Top10F"

/clone_lib="S21SNU520s1"

/note="Organ: Stomach; Vector: pTZ18Rpl; Site:1: EcoRI;

Site:2: NotI; The poly (A)+ RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then decapped

with tabacco acid pyrophosphatase (TAP). The decapped

intact mRNA was ligated with DNA-RNA linker including EcoR

I site by treatment of T4 RNA ligase and the first strand

cDNA was synthesized from oligo dt-selected mRNA by

priming with dt-tailed vector. The dt-tailed vector was

adjusted to have about 60nt. The cDNA vector was

circularized with E. coli DNA ligase after digestion of

EcoRI which site is also included in vector. An RNA strand

converted to a DNA strand by Okayama-Berg method. The

obtained cDNA vectors were used for transformation of

competent cells E. coli Top10F' by electroporation method.

The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dt)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method.

BASE COUNT 109 a 159 c 185 g 121 t
 ORIGIN

Alignment Scores:

Pred. No.: 2.72e-29 Length: 574
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

SEQ5 (1-59) x BM829991 (1-574)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 |||||
 Db 343 CAGAAATGAGGACTTGGGCCCGCCAGCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 402
 |||||
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
 |||||
 Db 403 GAGGACGATGACATGGGGGACCTGGTGATCTGAGGAGTATCTGGTACCCACGACGGC 462
 |||||
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
 |||||
 Db 463 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGTCTCCACACGACGACCGC 519
 |||||

RESULT 8

CB853376/c

LOCUS

DEFINITION

UI-CF-FN0-agd-h-18-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone

UI-CF-FN0-agd-h-18-0-UI 3', mRNA sequence.

ACCESSION CB853376

VERSION CB853376.1 GI:30048540

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 691)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

PUBMED

COMMENT

Contact: McCray, PB

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

Cloning Distribution by: Dr. M. Bento Soares, University of Iowa

Genetics (www.resgen.com) or from Open Biosystems

(www.openbiosystems.com).

Seq primer: M13 FORWARD
POLYA-No.

FEATURES

source

Location/Qualifiers

```
1. .691
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FNO"
/lab_host="Human Lung Epithelial cells"
/issue_type="PH10B (Life Technologies) (T1 phase resistant)"
/clone_lib="UI-CF-FNO"
/notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: Ecor I; Site_2: Not I; UI-CF-FNO is a subcloned cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DUL) The library was subcloned according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bent-soares@uiowa.edu
TAG_SEQ=None found"
```

BASE COUNT 140 a 224 c 191 g 131 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: 3,32e-29 Length: 691
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

SEQ05 (1-59) x CB853376 (1-691)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 317 CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCG 258
QY 21 GluAspAspMetGlyAspLeuValAlaGluTyrLeuValProGlnGlnGly 40
|||||
Db 257 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 198
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
|||||
Db 197 TTCTTCTGCCAGACCTGCCCGGGCGCTGGGGCATGCTCCACACAGGACCGC 141

RESULT 9
CA489799
LOCUS
DEFINITION AGENCOURT_10810428 MAPcL Homo sapiens cDNA clone IMAGE:6722285 5',
mRNA sequence.
ACCESSION CA489799
VERSION
KEYWORDS EST. GI:24952590
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 849)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL4284 row: m column: 05
High quality sequence stop: 539.

FEATURES

Location/Qualifiers

source

```
1. .849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722285"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1", LNCap"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/notes="Vector: pCMV-SPORT6; Site_1: EcorV; Site_2: Not I; Subcloned with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan. Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."
```

BASE COUNT 179 a 272 c 243 g 155 t
ORIGIN

Alignment Scores:
Pred. No.: 4,16e-29 Length: 849
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

SEQ05 (1-59) x CA489799 (1-849)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 185 CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCG 244
QY 21 GluAspAspMetGlyAspLeuValAlaGluTyrLeuValProGlnGlnGly 40
|||||
Db 245 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 304
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
|||||
Db 305 TTCTTCTGCCAGACCTGCCCGGGCGCTGGGGCATGCTCCACACAGGACCGC 361

RESULT 10

CD516283

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 852)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

CDNA Library Preparation: Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDAM492 row: j column: 17

High quality sequence stop: 662.

Location/Qualifiers

source

```

1..852
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="30408904"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_181"
/Note="Vector: PCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."
BASE COUNT 159 a 288 c 247 g 156 t 2 others
ORIGIN

```

Alignment Scores:

```

Pred. No.: 4.17e-29 Length: 852
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

```

SEQ5 (1-59) x CD516283 (1-852)

```

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
DB 95 CAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 154
|||||
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
|||||
DB 155 GAGGACGATGACATGGGGGAGCTGCTGGTGTCTGAGGAGTATCTGTACCCAGCAGGCGC 214
|||||
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
DB 215 TTCTTCTGTCAGACCTGCGCCGGCGCTGGGGCGATGGTCCACACGAGCAGCGC 271
|||||

```

RESULT 11

```

CA455074
LOCUS CA455074 885 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT_10735919 MAPcL Homo sapiens cDNA clone IMAGE:6722585 5',
mRNA sequence.
ACCESSION CA455074
VERSION CA455074.1 GI:24905427
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE 1 (bases 1 to 885)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14285 row: i column: 17
High quality sequence stop: 717.
Location/Qualifiers
1..885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="6722585"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1"

```

FEATURES

source

```

1..885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="6722585"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1"

```

```

' LNCap"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/Note="Vector: PCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bunkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."
BASE COUNT 178 a 297 c 261 g 149 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 4.35e-29 Length: 885
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

```

SEQ5 (1-59) x CA455074 (1-885)

```

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
DB 1 CAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 60
|||||
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
|||||
DB 61 GAGGACGATGACATGGGGGAGCTGCTGGTGTCTGAGGAGTATCTGTACCCAGCAGGCGC 120
|||||
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
DB 121 TTCTTCTGTCAGACCTGCGCCGGCGCTGGGGCGATGGTCCACACGAGCAGCGC 177
|||||

```

RESULT 12

```

CA455141
LOCUS CA455141 893 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT_10735980 MAPcL Homo sapiens cDNA clone IMAGE:6722663 5',
mRNA sequence.
ACCESSION CA455141
VERSION CA455141.1 GI:24905561
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE 1 (bases 1 to 893)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14285 row: 1 column: 23
High quality sequence stop: 682.
Location/Qualifiers
1..893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="6722663"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1"

```

FEATURES

source

```

1..893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="6722663"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1"

```

/note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

BASE COUNT 175 a 302 c 261 g 155 t
ORIGIN

Alignment Scores:
Pred. No.: 4,39e-29 Length: 893
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

SEQ5 (1-59) x CA455141 (1-893)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 67 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTCTG 126
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 127 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 186
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
Db 187 TTCTTCTGTCAGACCTGCCCCGGCGCTGGGGGATGTCACACAGGACCCGC 243

RESULT 13

CA454131

LOCUS CD558386 894 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT_10738550 MAPcL Homo sapiens cDNA clone IMAGE:6718792 5', mRNA sequence.

ACCESSION CA454131

VERSION CA454131.1 GI:24903569

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 894)

NIH-MGC http://mgi.nci.nih.gov/

Unpublished

National Institutes of Health, Mammalian Gene Collection (MGC)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Kristi A. Eglund, Ira Pastan

CDNA Library Preparation: Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM14275 row: k column: 16

High quality sequence stop: 636.

Location/Qualifiers

1. .894

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6718792"

/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HTERT-HME1, INCap"

/lab_host="EMDH10B"

/clone_lib="MAPcL"

/note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;

Subtracted with brain, liver, lung, kidney and muscle.

Directionally cloned. Priming method: oligo-dT. Average

insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

BASE COUNT 180 a 303 c 260 g 151 t
ORIGIN

Alignment Scores:
Pred. No.: 4,4e-29 Length: 894
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

SEQ5 (1-59) x CA454131 (1-894)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 24 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTCTG 83
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 84 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 143
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
Db 144 TTCTTCTGTCAGACCTGCCCCGGCGCTGGGGGATGTCACACAGGACCCGC 200

RESULT 14

CD558386

LOCUS CD558386

DEFINITION AGENCOURT_14476827 NIH_MGC_181 Homo sapiens cDNA clone IMAGE:30396412 5', mRNA sequence.

ACCESSION CD558386

VERSION CD558386.1 GI:31584454

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 899)

NIH-MGC http://mgi.nci.nih.gov/

Unpublished

National Institutes of Health, Mammalian Gene Collection (MGC)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

CDNA Library Preparation: Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDAM478 row: b column: 05

High quality sequence start: 47

High quality sequence stop: 443.

FEATURES

source

1. .899

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30396412"

/tissue_type="White Matter"

/dev_stage="Unknown"

/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"

/clone_lib="NIH_MGC_181"

/note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV (destroyed); Library is oligo-dT primed and directionally

cloned (EcoRV site is destroyed upon cloning). Average insert size 1.42 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."

BASE COUNT 165 a 338 c 237 g 159 t

ORIGIN

Alignment Scores:

Pred. No.: 4.42e-29 Length: 899
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

SEQ5 (1-59) x CD558386 (1-899)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 74 CAGAAATGAGGACTTGGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 133
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
|||||
Db 134 GAGGACGATGACATGGGGGAGCTGGTGGATGCTGAGGAGTATCTGGTACCCACGACGAGGC 193
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
|||||
Db 194 TTCTTCTGCCAGACCTTGGCCGGCGCTGGGGCGATGGTCCACACGACGACCGC 250

RESULT 15

BU845369

LOCUS

BU845369 916 bp mRNA linear EST 16-OCT-2002
AGENCOURT_10416348 NIH_MGC_109 Homo sapiens cDNA clone
IMAGE:6578980 5', mRNA sequence.

ACCESSION

BU845369.1 GI:24029810

VERSION

EST.

KEYWORDS

Source

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 916)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: L1CM2781 row: j column: 04
High quality sequence stop: 654.

FEATURES

source

1, 916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6578980"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_109"

/note="Organ: ovary; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 177 a 288 c 274 g 177 t

ORIGIN

Alignment Scores:

Pred. No.: 4.51e-29 Length: 916
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

SEQ5 (1-59) x BU845369 (1-916)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 343 CAGAAATGAGGACTTGGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 402
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
|||||
Db 403 GAGGACGATGACATGGGGGAGCTGGTGGATGCTGAGGAGTATCTGGTACCCACGACGAGGC 462
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
|||||
Db 463 TTCTTCTGCCAGACCTTGGCCGGCGCTGGGGCGATGGTCCACACGACGACCGC 519

Search completed: October 16, 2003, 17:04:51

Job time : 463.761 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 11:08:07 : Search time 60.9885 Seconds
(without alignments)
2540.503 Million cell updates/sec

Title: SEQ5

Perfect score: 326

Sequence: 1 QNEDLGASPDLSTFYRSL.....GFFCPDPAPGAGGVHRRH 59

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_P/HOLLERAN80/runat_15102003_131915_20662/app_query.fasta_1.4685
-DB=PublishedApplications_NA -OFF=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR_SCORE=ext -THR_MAX=100
-MAXLEN=2000000000 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=1 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=0 -YGAPOP=10 -XGAPOP=0.5
-YGAPOP=6 -YGAPOP=7 -YGAPOP=10 -YGAPOP=10 -YGAPOP=6 -DELEX=7

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	326	100.0	1115	14	US-10-102-806-165 Sequence 165, App

2	326	100.0	1713	12	US-10-378-393-14	Sequence 14, Appl
3	326	100.0	1755	10	US-09-930-125-6	Sequence 6, Appl
4	326	100.0	1767	10	US-09-930-125-4	Sequence 4, Appl
5	326	100.0	1773	10	US-09-930-125-7	Sequence 7, Appl
6	326	100.0	1806	10	US-09-930-125-5	Sequence 5, Appl
7	326	100.0	2411	12	US-10-378-393-10	Sequence 10, Appl
8	326	100.0	3765	12	US-10-207-498-5	Sequence 5, Appl
9	326	100.0	3768	9	US-09-811-123-8	Sequence 8, Appl
10	326	100.0	3768	9	US-09-811-115-2	Sequence 2, Appl
11	326	100.0	3768	10	US-09-854-356-9	Sequence 9, Appl
12	326	100.0	3768	10	US-09-930-125-1	Sequence 1, Appl
13	326	100.0	3768	12	US-10-313-644-1	Sequence 1, Appl
14	326	100.0	4473	11	US-09-441-411-5	Sequence 5, Appl
15	326	100.0	4473	12	US-10-101-510-81	Sequence 81, Appl
16	326	100.0	4473	14	US-10-146-473-32	Sequence 32, Appl
17	326	100.0	4473	14	US-10-207-655-44	Sequence 44, Appl
18	326	100.0	4530	10	US-09-877-177-11	Sequence 11, Appl
19	326	100.0	4530	12	US-10-007-926A-119	Sequence 119, App
20	326	100.0	4530	12	US-10-101-510-124	Sequence 124, App
21	326	100.0	4530	12	US-10-338-730-1	Sequence 1, Appl
22	326	100.0	4530	14	US-10-177-293-125	Sequence 125, App
23	326	100.0	4543	10	US-09-769-508-1	Sequence 1, Appl
24	326	100.0	4606	12	US-09-971-392-70	Sequence 70, Appl
25	326	100.0	4642	14	US-10-198-846-10896	Sequence 10896, A
26	326	100.0	9274	9	US-09-811-123-7	Sequence 7, Appl
27	326	100.0	9274	9	US-09-811-115-1	Sequence 1, Appl
28	279	85.6	3955	10	US-09-870-759-117	Sequence 117, App
29	279	85.6	3955	10	US-09-854-356-10	Sequence 10, Appl
30	279	85.6	3955	12	US-09-751-708A-117	Sequence 117, App
31	269	82.5	3771	10	US-09-854-356-11	Sequence 11, Appl
32	154	47.2	249	10	US-09-960-352-4186	Sequence 4186, Ap
33	138	42.3	3633	9	US-09-725-433-1	Sequence 1, Appl
34	138	42.3	5264	10	US-09-920-300A-1731	Sequence 1731, Ap
35	138	42.3	5264	12	US-10-099-926-1731	Sequence 1731, Ap
36	138	42.3	5264	13	US-10-033-528-1731	Sequence 137, Ap
37	138	42.3	5532	12	US-10-007-926A-137	Sequence 95, Appl
38	138	42.3	5532	12	US-10-101-510-95	Sequence 95, Appl
39	138	42.3	10058	10	US-09-874-298-98	Sequence 98, Appl
40	138	42.3	197496	10	US-09-977-177-10	Sequence 10, Appl
41	130	39.9	1935	12	US-10-259-451-1	Sequence 1, Appl
42	108.5	33.3	2437	11	US-09-158-722-3	Sequence 3, Appl
43	108.5	33.3	5484	10	US-09-940-101-1	Sequence 1, Appl
44	108.5	33.3	5484	14	US-10-207-655-46	Sequence 46, Appl
45	98	30.1	450	12	US-10-007-926A-136	Sequence 136, App

ALIGNMENTS

RESULT 1

US-10-102-806-165
; Sequence 165, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103p1c1
; CURRENT APPLICATION NUMBER: US/10/102,806
; PRIOR FILING DATE: 2002-03-22
; PRIOR FILING DATE: 2001-08-10
; PRIOR FILING DATE: 2000-03-08
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 165
; LENGTH: 1115
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (390)

```
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-165

Alignment Scores:
Pred. No.: 1.23e-38 Length: 1115
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

SEQ5 (1-59) x US-10-102-806-165 (1-1115)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 455 CAGAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCAGTCTG 514
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 515 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGC 574
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 575 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGTGTCCACACAGGCACCGC 631

RESULT 2
US-10-378-393-14
; Sequence 14, Application US/10378393
; Publication No. US20030182668A1
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
; FILE OF INVENTION: TYROSINE KINASE RECEPTORS
; CURRENT APPLICATION NUMBER: US/10/378,393
; PRIOR FILING DATE: 2003-03-03
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-14

Alignment Scores:
Pred. No.: 2.05e-38 Length: 1713
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ5 (1-59) x US-10-378-393-14 (1-1713)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 913 CAGAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCAGTCTG 972
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 973 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGC 1032
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 1032 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGTGTCCACACAGGCACCGC 1122

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-930-125-6

Alignment Scores:
Pred. No.: 2.11e-38 Length: 1755
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-930-125-6 (1-1755)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 946 CAGATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCAGTCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 1006 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGC 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGTGTCCACACAGGCACCGC 1122

RESULT 4
US-09-930-125-4
; Sequence 4, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-4
```

Alignment Scores:

Pred. No.: 2,13e-38 Length: 1767
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservativeness: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-930-125-4 (1-1767)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 |||||
 Db 946 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005
 QY 21 GluAspAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
 |||||
 Db 1006 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTACCCAGCAGGCG 1065
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArg 59
 |||||
 Db 1066 TTCTTCTGTCCAGACCTGCCCGGCGCTGGGGCGATGGTCCACCAGCAGCGC 1122

RESULT 5

US-09-930-125-7

; Sequence 7, Application US/09930125

; Publication No. US20020193329A1

GENERAL INFORMATION:

; APPLICANT: Hand-Zimmerman, Susan
 ; APPLICANT: Cheever, Martin A.
 ; APPLICANT: Foy, Teresa M.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Vedvick, Thomas S.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS

; FILE REFERENCE: 210121.544

; CURRENT APPLICATION NUMBER: US/09/930,125

; CURRENT FILING DATE: 2001-08-14

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 7

; LENGTH: 1773

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-930-125-7

Alignment Scores:

Pred. No.: 2,14e-38 Length: 1773
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservativeness: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-930-125-7 (1-1773)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 |||||
 Db 970 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1029
 QY 21 GluAspAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
 |||||
 Db 1030 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTACCCAGCAGGCG 1089
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArg 59
 |||||
 Db 1090 TTCTTCTGTCCAGACCTGCCCGGCGCTGGGGCGATGGTCCACCAGCAGCGC 1146

RESULT 6

US-09-930-125-5

; Sequence 5, Application US/09930125

; Publication No. US20020193329A1

GENERAL INFORMATION:

; APPLICANT: Hand-Zimmerman, Susan
 ; APPLICANT: Cheever, Martin A.
 ; APPLICANT: Foy, Teresa M.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Vedvick, Thomas S.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS

; FILE REFERENCE: 210121.544

; CURRENT APPLICATION NUMBER: US/09/930,125

; CURRENT FILING DATE: 2001-08-14

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 5

; LENGTH: 1806

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-930-125-5

Alignment Scores:

Pred. No.: 2,19e-38 Length: 1806
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservativeness: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-930-125-5 (1-1806)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 |||||
 Db 1003 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1062
 QY 21 GluAspAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
 |||||
 Db 1063 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTACCCAGCAGGCG 1122
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArg 59
 |||||
 Db 1123 TTCTTCTGTCCAGACCTGCCCGGCGCTGGGGCGATGGTCCACCAGCAGCGC 1179

RESULT 7

US-10-378-393-10

; Sequence 10, Application US/10378393

; Publication No. US20030182668A1

GENERAL INFORMATION:

; APPLICANT: Bol, David K.
 ; APPLICANT: Carboni, Joan M.
 ; APPLICANT: Rowley, Ronald B.
 ; APPLICANT: Wong, Tai W.
 ; APPLICANT: Lee, Francis Y.

; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVAT

; FILE REFERENCE: D0254 NP

; CURRENT APPLICATION NUMBER: US/10/378,393

; CURRENT FILING DATE: 2003-03-03

; PRIOR APPLICATION NUMBER: US 60/360,889

; PRIOR FILING DATE: 2002-03-01

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 10

; LENGTH: 2411

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-378-393-10

Alignment Scores:

Pred. No.: 3,09e-38 Length: 2411
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservativeness: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

SEQ5 (1-59) x US-10-378-393-10 (1-2411)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 Db 1611 CAGAATGAGGACTTGGCCCGCCAGCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1670
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
 Db 1671 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGCGC 1730
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
 Db 1731 TTCTTCTGTCAGACCTGCCCCGGGCGCTGGGGGCATGTTCCACACAGCAGCCGC 1787

RESULT 8

US-10-207-498-5
 ; Sequence 5, Application US/10207498
 ; Publication No. US20030143568A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Elizabeth Singer
 ; APPLICANT: Ralf Landgraf
 ; APPLICANT: Dennis J. Slamon
 ; APPLICANT: David Eisenberg
 ; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
 ; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREGULIN AND HER3
 ; FILE REFERENCE: 30448-103-US-U1
 ; CURRENT APPLICATION NUMBER: US/10/207,498
 ; PRIOR FILING DATE: 2002-07-29
 ; PRIOR APPLICATION NUMBER: 60/308,431
 ; PRIOR FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 3765
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(3765)
 ; US-10-207-498-5

Alignment Scores:
 Pred. No.: 5,25e-38 Length: 3765
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

SEQ5 (1-59) x US-10-207-498-5 (1-3765)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 Db 2968 CAGAATGAGGACTTGGCCCGCCAGCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
 Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGCGC 3087
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
 Db 3088 TTCTTCTGTCAGACCTGCCCCGGGCGCTGGGGGCATGTTCCACACAGCAGCCGC 3144

RESULT 9

US-09-811-123-8
 ; Sequence 8, Application US/09811123
 ; Patent No. US20020001587A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sharon Erickson

; APPLICANT: Ralph Schwall
 ; APPLICANT: Mark Sliwowski
 ; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ETBB
 ; TITLE OF INVENTION: ANTIBODY-MAYTANSINOID CONJUGATES
 ; FILE REFERENCE: GENENT.073A2
 ; CURRENT APPLICATION NUMBER: US/09/811,123
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/238,327
 ; PRIOR FILING DATE: 2000-10-05
 ; PRIOR APPLICATION NUMBER: 09/602,530
 ; PRIOR FILING DATE: 2000-06-23
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 3768
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-811-123-8

Alignment Scores:

Pred. No.: 5,26e-38 Length: 3768
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

SEQ5 (1-59) x US-09-811-123-8 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 Db 2968 CAGAATGAGGACTTGGCCCGCCAGCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
 Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGCGC 3087
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
 Db 3088 TTCTTCTGTCAGACCTGCCCCGGGCGCTGGGGGCATGTTCCACACAGCAGCCGC 3144

RESULT 10

US-09-811-115-2
 ; Sequence 2, Application US/09811115
 ; Patent No. US20020035736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Erickson, Sharon
 ; APPLICANT: Schwall, Ralph
 ; APPLICANT: King, Kathleen
 ; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
 ; FILE REFERENCE: GENENT.034A
 ; CURRENT APPLICATION NUMBER: US/09/811,115
 ; CURRENT FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/189,844
 ; PRIOR FILING DATE: 2000-03-16
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 3768
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-811-115-2

Alignment Scores:
 Pred. No.: 5,26e-38 Length: 3768
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

SEQ5 (1-59) x US-09-811-115-2 (1-3768)

```

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCACGACGAGGC 3087

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCTTCTGTCACAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACGACGACGCC 3144

RESULT 11
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.: 5,26e-38 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-854-356-9 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCACGACGAGGC 3087

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCTTCTGTCACAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACGACGACGCC 3144

RESULT 12
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US2002019329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3765)
US-09-930-125-1

Alignment Scores:
Pred. No.: 5,26e-38 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-930-125-1 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCACGACGAGGC 3087

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCTTCTGTCACAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACGACGACGCC 3144

RESULT 13
US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313,644
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:

```

```

Db 3088 TTCTTCTGTCACAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACGACGACGCC 3144

RESULT 12
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US2002019329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3765)
US-09-930-125-1

Alignment Scores:
Pred. No.: 5,26e-38 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-930-125-1 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCACGACGAGGC 3087

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCTTCTGTCACAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACGACGACGCC 3144

RESULT 13
US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313,644
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:

```

```
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-313-644-1

Alignment Scores:
Pred. No.: 5.26e-38 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ5 (1-59) x US-10-313-644-1 (1-3768)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGANTGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTGTTACCCAGCAGGC 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGGCGATGTCACACACAGGCACCGC 3144

RESULT 14
US-09-441-411-5
; Sequence 5, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-5

Alignment Scores:
Pred. No.: 6.45e-38 Length: 4473
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

SEQ5 (1-59) x US-09-441-411-5 (1-4473)
QY 1 GluAspGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGANTGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTGTTACCCAGCAGGC 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3262 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGGCGATGTCACACACAGGCACCGC 3318

RESULT 15
US-10-101-510-81
; Sequence 81, Application US/10101510
; Publication No. US20030148295A1
```

```
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-81

Alignment Scores:
Pred. No.: 6.45e-38 Length: 4473
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ5 (1-59) x US-10-101-510-81 (1-4473)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGANTGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTGTTACCCAGCAGGC 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3262 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGGCGATGTCACACACAGGCACCGC 3318

Search completed: October 17, 2003, 03:54:41
Job time : 65.9885 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:55:19 ; Search time 14.9563 Seconds
(without alignments)
1741.185 Million cell updates/sec

Title: SEQ5

Perfect score: 326

Sequence: 1 QNEDLGASPILDSTFYRSL.....GFFCPDPAPGAGGVHRRH 59

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O/cgn2_1/USPTO_spool_P/HOLLERAN480/runat_15102003_131913_20570/app_query.fasta_1.4685
-DB-Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 @cgn_1_1_403 @runat_15102003_131913_20570 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	326	100.0	3768	2	US-08-625-101-1
2	326	100.0	3768	2	US-08-356-786-1
3	326	100.0	4473	3	US-09-048-804-1
4	326	100.0	4473	3	US-09-056-105-26
5	326	100.0	4530	1	US-08-229-515A-9
6	326	100.0	4530	1	US-08-645-865-9
7	326	100.0	4530	4	US-09-167-322-4
8	326	100.0	4530	4	US-09-527-487-1
9	326	100.0	4530	4	US-09-877-177A-11
10	279	85.6	3955	1	US-08-229-515A-14
11	279	85.6	3955	1	US-08-645-865-14
12	138	42.3	5532	2	US-08-475-035-3

13	138	42.3	5532	4	US-09-676-610B-17	Sequence 17, Appl
14	138	42.3	197496	4	US-09-877-177A-10	Sequence 10, Appl
15	136	41.7	410	3	US-08-604-991-11	Sequence 11, Appl
16	136	41.7	410	3	US-09-363-639-11	Sequence 11, Appl
17	136	41.7	576	1	US-08-783-275-3	Sequence 3, Appl
18	136	41.7	576	1	US-08-727-708-3	Sequence 3, Appl
19	136	41.7	576	2	US-08-766-677-1	Sequence 1, Appl
20	136	41.7	576	2	US-08-843-951-1	Sequence 1, Appl
21	108.5	33.3	2437	2	US-08-456-647B-3	Sequence 3, Appl
22	108.5	33.3	2437	2	US-08-237-401A-3	Sequence 3, Appl
23	108.5	33.3	3321	1	US-08-484-438-5	Sequence 5, Appl
24	108.5	33.3	5484	3	US-09-632-580A-3	Sequence 3, Appl
25	108.5	33.3	5501	1	US-08-484-438-1	Sequence 3, Appl
26	108.5	33.3	5555	1	US-08-484-438-3	Sequence 6, Appl
27	66	20.2	980	4	US-09-218-467B-6	Sequence 6, Appl
28	66	20.2	1209	4	US-09-170-496D-51	Sequence 51, Appl
29	66	20.2	1209	4	US-09-170-496D-191	Sequence 191, App
30	66	20.2	1269	3	US-09-224-426-1	Sequence 1, Appl
31	66	20.2	1269	3	US-09-478-601-1	Sequence 1, Appl
32	66	20.2	1269	3	US-09-478-602-1	Sequence 1, Appl
33	66	20.2	1316	3	US-08-602-809-1	Sequence 1, Appl
34	66	20.2	1316	5	PCT-US95-16472-1	Sequence 1, Appl
35	66	20.2	1385	3	US-08-984-288-1	Sequence 1, Appl
36	66	20.2	3488	4	US-09-218-467B-1	Sequence 1, Appl
37	65	19.9	1173	2	US-08-614-156B-4	Sequence 4, Appl
38	65	19.9	1173	3	US-09-213-864-1	Sequence 1, Appl
39	64	19.6	6462	4	US-09-792-594-10	Sequence 10, Appl
40	62.5	19.2	975	4	US-09-252-991A-12244	Sequence 12244, A
41	62.5	19.2	1518	4	US-09-252-991A-12104	Sequence 12104, A
42	62.5	19.2	2040	4	US-09-252-991A-12316	Sequence 12316, A
43	62	19.0	312	4	US-09-252-991A-16553	Sequence 16553, A
44	62	19.0	501	4	US-09-252-991A-10705	Sequence 10705, A
45	62	19.0	999	4	US-09-252-991A-15955	Sequence 15955, A

ALIGNMENTS

RESULT 1

US-08-625-101-1

; Sequence 1, Application US/08625101

; Patent No. 5863445

; GENERAL INFORMATION:

; APPLICANT: Cheever, Martin A.

; APPLICANT: Disis, Mary L.

; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE

; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION

; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu

; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/625,101

; FILING DATE: 01-APR-1996

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Sharkey, Richard G.

; REGISTRATION NUMBER: 32,629

; REFERENCE/DOCKET NUMBER: 920010.448C7

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 3768 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 3,24e-32 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

SEQ5 (1-59) x US-08-625-101-1 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 2968 CAGATGAGGAGCTTGGGCCAGCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40
DB 3028 GAGGACGATGACATGGGGACCTGTGGATGCTGAGGAGTATCTGTATCCCCCAGCAGGCG 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
DB 3088 TTCTTCTGTCAGACCCCTGCCCGGGCGCTGGGGGCATGGTCCACACAGGACCGC 3144

RESULT 2

US-08-356-786-1
Sequence 1, Application US/08356786
Patent No. 5877305

GENERAL INFORMATION:
APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
TITLE OF INVENTION: Marker
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3768 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3768
OTHER INFORMATION: /note= "product = "cerB-b2"
US-08-356-786-1

Alignment Scores:
Pred. No.: 3,24e-32 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

SEQ5 (1-59) x US-08-356-786-1 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 2968 CAGATGAGGAGCTTGGGCCAGCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40
DB 3028 GAGGACGATGACATGGGGACCTGTGGATGCTGAGGAGTATCTGTATCCCCCAGCAGGCG 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
DB 3088 TTCTTCTGTCAGACCCCTGCCCGGGCGCTGGGGGCATGGTCCACACAGGACCGC 3144

RESULT 3

US-09-048-804-1
Sequence 1, Application US/09048804
Patent No. 5968748

GENERAL INFORMATION:
APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
TITLE OF INVENTION: HUMAN HER-2 EXPRESSION
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748rls LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 Mb diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/048,804
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Paul K. Legaard
REGISTRATION NUMBER: 38,534
REFERENCE/DOCKET NUMBER: ISIS-2913
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4473 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
ANTI-SENSE: No

US-09-048-804-1

Alignment Scores:

Pred. No.: 3,96e-32 Length: 4473
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

SEQ5 (1-59) x US-09-048-804-1 (1-4473)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCACGACGCGC 3261
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
Db 3262 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACGACGACCGC 3318

RESULT 4

US-09-056-105-26

; Sequence 26, Application US/09056105

; Patent No. 6287569

; GENERAL INFORMATION:

; APPLICANT: KIPPS, THOMAS J.

; APPLICANT: WU, YUNQI

; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR

; FILE OF INVENTION: PROCESSING

; CURRENT APPLICATION NUMBER: US/09/056,105

; EARLIER FILING DATE: 1998-04-06

; EARLIER APPLICATION NUMBER: 60/043,467

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 26

; LENGTH: 4473

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-056-105-26

Alignment Scores:

Pred. No.: 3,96e-32 Length: 4473
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

SEQ5 (1-59) x US-09-056-105-26 (1-4473)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCACGACGCGC 3261
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
Db 3262 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACGACGACCGC 3318

RESULT 5

US-08-229-515A-9

; Sequence 9, Application US/08229515A

; Patent No. 5518885

; GENERAL INFORMATION:

; APPLICANT: RAZIUDDIN

; APPLICANT: SARKAR, FAZLUL H

; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08/229,515A

APPLICATION NUMBER: 19 APR 1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: PERRYMAN, DAVID G

REGISTRATION NUMBER: 33,438

REFERENCE/DOCKET NUMBER: 1414,608

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-688-0770

TELEFAX: 404-688-9880

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 4530 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-229-515A-9

Alignment Scores:

Pred. No.: 4.01e-32 Length: 4530
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

SEQ5 (1-59) x US-08-229-515A-9 (1-4530)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3118 CAGAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3177
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3178 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCACGACGCGC 3237
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
Db 3238 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACGACGACCGC 3294

RESULT 6

US-08-645-865-9

; Sequence 9, Application US/08645865

; Patent No. 5654406

; GENERAL INFORMATION:

; APPLICANT: RAZIUDDIN

; APPLICANT: SARKAR, FAZLUL H

; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN

; TITLE OF INVENTION: NEOPLASTIC DISEASE

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NEEDLE & ROSENBERG PC

; STREET: 127 Peachtree Street, Suite 1200

; CITY: Atlanta

; STATE: Georgia

; COUNTRY: usa

; ZIP: 30303


```
Db 3118 CAGAATGAGGACTTGGGGCCAGCCAGTCCCTTGGAGACACCTTCTACCGCTCACTGCTG 3177
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3178 GAGGACGATGACATGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3237
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3238 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGGTCCACACAGCAGCCGC 3294

RESULT 9
US-09-877-177A-11
; Sequence 11, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11

Alignment Scores:
Pred. No.: 4,01e-32 Length: 4530
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

SEQ5 (1-59) x US-09-877-177A-11 (1-4530)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3118 CAGAATGAGGACTTGGGGCCAGCCAGTCCCTTGGAGACACCTTCTACCGCTCACTGCTG 3177
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3178 GAGGACGATGACATGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3237
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3238 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGGTCCACACAGCAGCCGC 3294

RESULT 10
US-08-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A-14

Alignment Scores:
Pred. No.: 3,49e-26 Length: 3955
Score: 279.00 Matches: 50
Percent Similarity: 88.14% Conservative: 2
Best Local Similarity: 84.75% Mismatches: 7
Query Match: 85.58% Indels: 0
DB: 1 Gaps: 0

SEQ5 (1-59) x US-08-229-515A-14 (1-3955)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2999 CAGACGAGGACTGGGGCCATCCAGCCCCAGGACAGTACCTTACCGTCACTGCTG 3058
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3059 GAAGATGATGACATGGGTGACCTGGTAGACGCTGAAGATATCTGGTGCCTCCAGCAGGGA 3118
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3119 TTCTTCTCCCGGACCCCTACCCAGGCACCTGGGAGCACAGCCCATAGAAGGCACCGC 3175

RESULT 11
US-08-645-865-14
; Sequence 14, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
```

```
;
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-645-865-14

Alignment Scores:
Pred. No.: 3.49e-26 Length: 3955
Score: 279.00 Matches: 50
Percent Similarity: 88.14% Conservatives: 2
Best Local Similarity: 84.75% Mismatches: 7
Query Match: 85.58% Indels: 0
DB: 1 Gaps: 0

SEQ5 (1-59) x US-08-645-865-14 (1-3955)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2999 CAGAAGCAGGACTGGGCCCAATCCAGCCCATGAGACAGTACCTTCTACCGTTCCAGCTG 3058
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3059 GAAGATGATGACATGGTGGTACCTGGTAGACGCTGAAGAGTATCTGCTGCCCCAGCAGGGA 3118
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3119 TTCTTCCCGGACCTACCCAGGCACTGGGAGCAGACGCCCATAGAGGCCCGC 3175

RESULT 12
US-08-475-035-3
; Sequence 3, Application US/08475035
; Patent No. 5985553
; GENERAL INFORMATION:
; APPLICANT: KING, C. R.
; APPLICANT: KRAUS, MATTHIAS H.
; APPLICANT: AARONSON, STUART A.
; TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM
; TITLE OF INVENTION: EGF RECEPTOR GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, 127 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,035
; FILING DATE: 7 Jun 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.656
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 187...3816
US-08-475-035-3

Alignment Scores:
Pred. No.: 5.46e-08 Length: 5532
Score: 138.00 Matches: 23
Percent Similarity: 91.18% Conservatives: 8
Best Local Similarity: 67.65% Mismatches: 3
Query Match: 42.33% Indels: 0
DB: 2 Gaps: 0

SEQ5 (1-59) x US-08-475-035-3 (1-5532)
QY 9 SerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeu 28
Db 3157 AGTCCTACAGACTCCAACTTCTACCGTGCCTGATGATGAAGAAGACATGGACGCGT 3216
QY 29 ValAspAlaGluTyrLeuValProGlnGlnGlyPhePhe 42
Db 3217 GTGGATGGCAGCAGTACCTATCCACACAGCGGCTTCTTC 3258

RESULT 13
US-09-676-610B-17
; Sequence 17, Application US/09676610B
; Patent No. 6444465
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Jacqueline Wyatt
; APPLICANT: Susan M. Freiler
; TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
; FILE REFERENCE: RTS-0138
; CURRENT APPLICATION NUMBER: US/09/676,610B
; CURRENT FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 182
; SEQ ID NO 17
; LENGTH: 5532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (187)...(3819)
US-09-676-610B-17

Alignment Scores:
Pred. No.: 5.46e-08 Length: 5532
Score: 138.00 Matches: 23
Percent Similarity: 91.18% Conservatives: 8
Best Local Similarity: 67.65% Mismatches: 3
Query Match: 42.33% Indels: 0
DB: 4 Gaps: 0

SEQ5 (1-59) x US-09-676-610B-17 (1-5532)
QY 9 SerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeu 28
Db 3157 AGTCCTACAGACTCCAACTTCTACCGTGCCTGATGATGAAGAAGACATGGACGCGT 3216
QY 29 ValAspAlaGluTyrLeuValProGlnGlnGlyPhePhe 42
Db 3217 GTGGATGGCAGCAGTACCTATCCACACAGCGGCTTCTTC 3258

RESULT 14
US-09-877-177A-10
; Sequence 10, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
```

```

; MOLECULE TYPE: DNA (genomic)
US-08-604-991-11

Alignment Scores:
Pred. No.: 4.77e-09 Length: 410
Score: 136.00 Matches: 24
Percent Similarity: 88.24% Conservativeness: 6
Best Local Similarity: 70.59% Mismatches: 4
Query Match: 41.72% Indels: 0
DB: 3 Gaps: 0

SEQ5 (1-59) x US-08-604-991-11 (1-410)
QY 9 SerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeu 28
Db 225 AGCCCTACAGATTCCAAAGTTTATCGCACCTGATGGAGGAGGACATGGAGACATT 284
QY 29 ValAspAlaGluTyrLeuValProGlnGlnGlyPhephe 42
Db 285 GTGGATGCAGATGATGATCTTGTCCACACACAGGGCTTTTTC 326

Search completed: October 16, 2003, 17:18:38
Job time : 41.9563 secs

; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 197496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-10

Alignment Scores:
Pred. No.: 3.49e-06 Length: 197496
Score: 138.00 Matches: 23
Percent Similarity: 91.18% Conservativeness: 8
Best Local Similarity: 67.65% Mismatches: 3
Query Match: 42.33% Indels: 0
DB: 4 Gaps: 0

SEQ5 (1-59) x US-09-877-177A-10 (1-197496)
QY 9 SerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeu 28
Db 191318 AGTCCTACAGACTCCAACTTCTACCGTGCCCTGATGGATGAAGAAGACATGGACGCGTG 191377
QY 29 ValAspAlaGluTyrLeuValProGlnGlnGlyPhephe 42
Db 191378 GTGGATGCCGAGGATGATCTATCCACAGCAGGGCTTCTTC 191419

RESULT 15
US-08-604-991-11
; Sequence 11, Application US/08604991
; Patent No. 6001565
; GENERAL INFORMATION:
; APPLICANT: Fox, Andrew J.
; APPLICANT: Jones, Dennis Mackay
; TITLE OF INVENTION: Detection and Speciation of
; TITLE OF INVENTION: Campylobacter
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSER: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/604,991
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/01967
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9318751.6
; FILING DATE: 09-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1581.0040000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-317-2540
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 56.0384 Seconds
(without alignments)
2842.104 Million cell updates/sec

Title: SEQ5

Perfect score: 326

Sequence: 1 QNEDIGPASPLDSTFYRSLL.....GFFCPDPAPGAGGMVHHRH 59

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cn2.1/USPTO_spool_p/HOLLERAN480/runat_15102003_131912_20526/app_query.fasta_1.4685
-DB=N_Geneseq_19jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGN 1.1 1758 @runat_15102003_131912_20526 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -FGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	326	100.0	1115	21	AAF21778	Human breast and o
2	326	100.0	1755	24	AAD32746	Human CDNA for the
3	326	100.0	1767	24	AAD32744	Human CDNA for the
4	326	100.0	1773	24	AAD32747	Human CDNA for the
5	326	100.0	1806	24	AAD32745	Human CDNA for the
6	326	100.0	3600	21	AAA89736	Human HER-2/neu co
7	326	100.0	3678	24	ABK86207	cDNA encoding huma
8	326	100.0	3768	17	AAT40237	HER-2/neu oncogene
9	326	100.0	3768	20	AAK01912	Human HER-2/neu on
10	326	100.0	3768	21	AAA09455	Human heregulin 2
11	326	100.0	3768	22	AAH23392	Human HER-2/neu pr
12	326	100.0	3768	24	ABZ35744	Human ERBB2 polyinu
13	326	100.0	3768	24	ABX09987	Human ERBB2 DNA fr
14	326	100.0	3768	24	AAD43935	Human HER-2 cDNA.
15	326	100.0	3768	24	AAD43986	Human HER2 antigen
16	326	100.0	3768	24	ABV78168	Human ERBB2 DNA SE
17	326	100.0	3768	24	AAD32743	Human HER-2/neu pr
18	326	100.0	3768	24	ABA92250	Human HER-2/neu CD
19	326	100.0	3768	24	ABK10730	Human HER-2/neu DN
20	326	100.0	3768	24	ABL91709	Human polynucleoti
21	326	100.0	3768	24	ABK114058	Human HER2 (Erbb2)
22	326	100.0	4299	14	AAQ46083	Sequence encoding
23	326	100.0	4472	21	AAA14812	cDNA encoding the
24	326	100.0	4473	19	ABQ76220	Human tumour antig
25	326	100.0	4473	20	AAZ31071	HER-2 nucleic acid
26	326	100.0	4473	24	ABZ34969	Human gene express
27	326	100.0	4473	24	ABZ38904	Human Her-2 DNA.
28	326	100.0	4530	16	AAT01585	Her-2/neu (ERBB2/c
29	326	100.0	4530	18	AAT71253	Human HER2 gene.
30	326	100.0	4530	21	AAZ60815	Nucleotide sequenc
31	326	100.0	4530	22	AAD19731	Human tyrosine kin
32	326	100.0	4530	24	ABZ35012	Human gene express
33	326	100.0	4530	24	ABV94128	Breast carcinoma r
34	326	100.0	4530	24	ABN85585	Human HER2-neu SEQ
35	326	100.0	4530	24	ABK83918	Human CDNA differe
36	326	100.0	4530	25	ACC50139	Breast cancer asso
37	326	100.0	4530	25	ABQ83856	Human Her2/Neu enc
38	326	100.0	9274	22	AAF24297	HER2 transgene pla
39	326	100.0	9274	24	AAD43934	HER-2 transgene pl
40	326	100.0	9274	24	ABK14057	Human HER2 (Erbb2)
41	279	85.6	3955	16	AAT01590	Rat neu promoter.
42	279	85.6	3955	21	AAA89753	Rat HER-2/neu prot
43	269	82.5	2763	24	ABA92252	Mouse Her-2/neu ex
44	269	82.5	2781	24	ABA92253	Her-2/neu extracel
45	269	82.5	3771	21	AAA89737	Mouse Her-2/neu CD

ALIGNMENTS

RESULT 1

AAF21778

ID AAF21778 standard; DNA; 1115 BP.

XX

AC AAF21778;

DT 27-MAR-2001 (first entry)

XX Human breast and ovarian cancer associated antigen gene SEQ ID 165.

DE Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;

KW cardiovascular disorder; wound healing; neurological disease; ds.
 XX Homo sapiens.
 OS
 PN W0200055173-Al.
 XX
 PD 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US05881.
 XX
 XX 12-MAR-1999; 99US-0124270.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2000-611515/58.
 DR P-PSDB; AAB56875.
 XX
 XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases
 XX
 XX Claim 1; Page 604; 1299pp; English.
 PS
 XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;
 CC antibacterial; antifungal; antiparasitic and cardiant activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and agonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemias; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 XX
 XX Sequence 1115 BP; 210 A; 336 C; 338 G; 222 T; 9 other;
 SQ

Alignment Scores:
 Pred. No.: 1.4e-32 Length: 1115
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

SEQ5 (1-59) x AAF21778 (1-1115)
 Qy 1 GLuAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
 Db 455 CAGNATGAGGACTGGGCCCCAGCAGTCCCTTGACAGACACCTTCTACCGCTCAGTGTG 514
 Qy 21 GLuAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
 Db 515 GAGGACCATGACATGGGGGACCTGGTGGATGCTGAGAGATATCTGGTATCCCGACGAGGC 574
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
 Db 575 TTCTTCGTGTCACACCTGCCCGGGGCTGGGGGCGATGTCACACACGACGACCGC 631
 RESULT 2
 AAD32746
 ID AAD32746 standard; cDNA; 1755 BP.

XX AAD32746;
 AC
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Human cDNA for the clone HICD_native_coding_region.
 XX
 KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 1..1755
 FT /tag= a
 FT /product= "Human protein encoded by cDNA for the clone
 FT HICD_native_coding_region"
 FT /transl_except= (pos:1741..1752, aa:Leu-Glu)
 FT /note= "CDS does not include stop codon"
 FT /partial
 XX
 XX W0200214503-A2.
 XX
 XX 21-FEB-2002.
 PD
 XX
 XX 14-AUG-2001; 2001WO-US41733.
 XX
 XX 14-AUG-2000; 2000US-225152P.
 PR
 XX 28-SEP-2000; 2000US-236428P.
 PR
 XX 21-FEB-2001; 2001US-270520P.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kaloş MD;
 PI McNeill PD, Vedvick TS;
 PI
 XX WPI; 2002-280758/32.
 DR P-PSDB; AAE20483.
 XX
 XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,
 XX prevention and diagnosis of cancer, preferably breast cancer
 XX
 XX Example 5; Page 118-119; 129pp; English.
 XX
 XX The invention relates to an isolated Her-2/Neu polypeptide composition
 XX effective for eliciting an immune response. The invention is useful for
 XX eliciting an immune response in a patient, where the patient is human
 XX leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
 XX The composition is useful for the therapy and diagnosis of cancer,
 XX and preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
 XX and other compositions for the diagnosis, prevention and treatment of
 XX human malignancies, for stimulating and/or expanding T cells specific for
 XX Her-2/Neu polypeptide and for inhibiting the development of cancer in a
 XX patient. The invention is useful for stimulating a T cell response in a
 XX human patient, as probe or primer for nucleic acid hybridisation, to
 XX selectively form duplex molecules with complementary stretches of the
 XX entire Her-2/Neu gene or gene fragments of interest, to isolate a full
 XX length gene from a suitable library, and to direct expression of a
 XX polypeptide in appropriate host cells. The composition is useful in
 XX prophylactic or therapeutic applications and for the treatment of cancer,
 XX preferably for the immunotherapy of breast cancer and other Her-2/Neu-
 XX associated malignancies. The invention is useful in gene therapy. The
 XX present sequence is human cDNA for the clone HICD_native_coding_region.
 XX
 XX Sequence 1755 BP; 376 A; 517 C; 531 G; 331 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 2.4e-32 Length: 1755
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

```
SEQ5 (1-59) x AAD32746 (1-1755)
QY      1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
      |||||||
Db      946 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005
      |||||||
QY      21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
      |||||||
Db      1006 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTACCCAGCAGGCG 1065
      |||||||
QY      41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
      |||||||
Db      1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGGTCCACACAGGACCGCG 1122
      |||||||

RESULT 3
AAD32744
ID      AAD32744 standard; cDNA; 1767 BP.
AC      AAD32744;
DT      01-JUL-2002 (first entry)
DE      Human cDNA for the clone HICD_CT_His_coding_region.
KW      Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
      human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
OS      Homo sapiens.
FH      Key
FT      CDS
      Location/Qualifiers
      1..1764
      /*tag= a
      /product= "Human protein encoded by cDNA for the clone
      HICD_CT_His_coding_region"

WO200214503-A2.
21-FEB-2002.
14-AUG-2001; 2001WO-US41733.
14-AUG-2000; 2000US-225152P.
28-SEP-2000; 2000US-236428P.
21-FEB-2001; 2001US-270520P.
(CORI-) CORIXA CORP.
Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
Mcneill PD, Vedvick TS;
WPI; 2002-280758/32.
P-PSDB; AAE20481.

Novel isolated Her-2/Neu polypeptide composition useful for therapy,
prevention and diagnosis of cancer, preferably breast cancer -
Example 5; Page 117-118; 129pp; English.

The invention relates to an isolated Her-2/Neu polypeptide composition
effective for eliciting an immune response. The invention is useful for
eliciting an immune response in a patient, where the patient is human
leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
The composition is useful for the therapy and diagnosis of cancer,
preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
and other compositions for the diagnosis, prevention and treatment of
human malignancies, for stimulating and/or expanding T cells specific for
Her-2/Neu polypeptide and for inhibiting the development of cancer in a
patient. The invention is useful for stimulating a T cell response in a
human patient, as probe or primer for nucleic acid hybridisation, to
selectively form duplex molecules with complementary stretches of the
entire Her-2/Neu gene or gene fragments of interest, to isolate a full
length gene from a suitable library, and to direct expression of a
```

```
CC polypeptide in appropriate host cells. The composition is useful in
CC prophylactic or therapeutic applications and for the treatment of cancer,
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
CC associated malignancies. The invention is useful in gene therapy. The
CC present sequence is human cDNA for the clone HICD_CT_His_coding_region.
XX
SQ      Sequence 1767 BP; 381 A; 521 C; 529 G; 336 T; 0 other;

Alignment Scores:
Pred. No.:      2.42e-32      Length:      1767
Score:          326.00      Matches:      59
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              24      Gaps:      0

SEQ5 (1-59) x AAD32744 (1-1767)
QY      1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
      |||||||
Db      946 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005
      |||||||
QY      21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
      |||||||
Db      1006 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTACCCAGCAGGCG 1065
      |||||||
QY      41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
      |||||||
Db      1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGGTCCACACAGGACCGCG 1122
      |||||||

RESULT 4
AAD32747
ID      AAD32747 standard; cDNA; 1773 BP.
XX
AC      AAD32747;
DT      01-JUL-2002 (first entry)
XX
DE      Human cDNA for the clone HICD_in_ppDM_coding_sequence.
XX
KW      Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
      human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
OS      Homo sapiens.
XX
FH      Key
FT      CDS
      Location/Qualifiers
      1..1770
      /*tag= a
      /product= "Human protein encoded by cDNA for the clone
      HICD_in_ppDM_coding_sequence"

WO200214503-A2.
21-FEB-2002.
14-AUG-2001; 2001WO-US41733.
14-AUG-2000; 2000US-225152P.
28-SEP-2000; 2000US-236428P.
21-FEB-2001; 2001US-270520P.
(CORI-) CORIXA CORP.
Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
Mcneill PD, Vedvick TS;
WPI; 2002-280758/32.
P-PSDB; AAE20481.

Novel isolated Her-2/Neu polypeptide composition useful for therapy,
prevention and diagnosis of cancer, preferably breast cancer -
Example 5; Page 119; 129pp; English.
```

```

XX CC The invention relates to an isolated Her-2/Neu polypeptide composition
CC effective for eliciting an immune response. The invention is useful for
CC eliciting an immune response in a patient, where the patient is human
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
CC The composition is useful for the therapy and diagnosis of cancer,
CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
CC and other compositions for stimulating and/or expanding T cells specific for
CC human malignancies, for stimulating and/or inhibiting the development of a
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
CC patient. The invention is useful for stimulating a T cell response in a
CC human patient, as probe or primer for nucleic acid hybridisation, to
CC selectively form duplex molecules with complementary stretches of the
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
CC length gene from a suitable library, and to direct expression of a
CC polypeptide in appropriate host cells. The composition is useful in
CC prophylactic or therapeutic applications and for the treatment of cancer,
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
CC associated malignancies. The invention is useful in gene therapy. The
CC present sequence is human cDNA for the clone HICD_in_ppdm_coding_
XX CC
SQ Sequence 1773 BP; 383 A; 528 C; 530 G; 332 T; 0 other;

Alignment Scores:
Pred. No.: 2,43e-32 Length: 1773
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ5 (1-59) x AAD32747 (1-1773)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 970 CAGAAATGAGGACTGTGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCAGTGCTG 1029
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 1030 GAGGACGATGACATGGGGACCTGGTGATGCTGAGAGTATCTGTATCCCGCAGCAGGC 1089
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 1090 TTCTTGTGTGACAGACCTGCCCGGGGCTGGGGGATGCTGCCACACAGCAGCCG 1146

RESULT 5
AAD32745
ID AAD32745 standard; cDNA; 1806 BP.
XX AC AAD32745;
XX XX
XX DT 01-JUL-2002 (first entry)
XX XX
DE Human cDNA for the clone HICD_plus_8_HIS.
XX XX
KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
XX XX
OS Homo sapiens.
XX XX
FH Key
FT CDS 1..1803
FT /tag= a
FT /product= "Human protein encoded by cDNA for the clone
FT Hicb_plus_8_HIS"
FT /transl_except= (pos:1543..1545, aa:Pro)
XX XX
PN WO200214503-A2.
XX XX
PD 21-FEB-2002.
XX XX
PF 14-AUG-2001; 2001WO-US41733.

```

```

XX PR 14-AUG-2000; 2000US-225152P.
PR 28-SEP-2000; 2000US-236428P.
PR 21-FEB-2001; 2001US-270520P.
XX PA (CORI-) CORIXA CORP.
XX PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
PI McNeill PD, Vedvick TS;
XX XX
XX WPI; 2002-280758/32.
XX P-PSDB; AAE20482.
XX PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,
XX prevention and diagnosis of cancer, preferably breast cancer -
XX PS Example 5; Page 118; 129pp; English.
XX XX
XX CC The invention relates to an isolated Her-2/Neu polypeptide composition
CC effective for eliciting an immune response. The invention is useful for
CC eliciting an immune response in a patient, where the patient is human
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
CC The composition is useful for the therapy and diagnosis of cancer,
CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
CC and other compositions for the diagnosis, prevention and treatment of
CC human malignancies, for stimulating and/or expanding T cells specific for
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
CC patient. The invention is useful for stimulating a T cell response in a
CC human patient, as probe or primer for nucleic acid hybridisation, to
CC selectively form duplex molecules with complementary stretches of the
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
CC length gene from a suitable library, and to direct expression of a
CC polypeptide in appropriate host cells. The composition is useful in
CC prophylactic or therapeutic applications and for the treatment of cancer,
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
CC associated malignancies. The invention is useful in gene therapy. The
CC present sequence is human cDNA for the clone HICD_plus_8_HIS.
XX CC
SQ Sequence 1806 BP; 391 A; 530 C; 544 G; 341 T; 0 other;

Alignment Scores:
Pred. No.: 2,48e-32 Length: 1806
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ5 (1-59) x AAD32745 (1-1806)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 1003 CAGAAATGAGGACTGTGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCAGTGCTG 1062
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 1063 GAGGACGATGACATGGGGACCTGGTGATGCTGAGAGTATCTGTATCCCGCAGCAGGC 1122
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 1123 TTCTTGTGTGACAGACCTGCCCGGGGCTGGGGGATGCTGCCACACAGCAGCCG 1179

RESULT 6
AAD32745
ID AAD32745 standard; DNA; 3600 BP.
XX AC AAD32745;
XX XX
XX DT 12-JAN-2001 (first entry)
XX XX
DE Human HER-2/neu coding sequence.
XX XX
PF Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;

```



```

Db      2878 CAGAATGAGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTACCGCTCAGTCTG 2937
QY      21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
      |||||
Db      2938 GAGGACGATGACATGGGGACCTGGTGGATCTGAGAGATATCTGTGATCCCGACGAGGC 2997
QY      41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
      |||||
Db      2998 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGTGTCCACACAGGCACCGC 3054

RESULT 8
AAT40739
ID      AAT40739 standard; cDNA; 3768 BP.
XX
AC      AAT40739;
XX
DT      01-JAN-1997 (first entry)
XX
DE      HER-2/neu oncogene.
XX
KW      HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
KW      breast cancer; ovary cancer; colon cancer; lung cancer;
KW      prostate cancer; genetic immunisation; tumour; vaccine; vector;
KW      ss.
OS      Homo sapiens.
FH      Key
FT      CDS
      1..3765
      /*tag= b
      /note= "nucleotides 2026-3765 (claim 1) code for
      HER-2/neu intracellular domain"
XX
PN      WO9630514-A1.
XX
PD      03-OCT-1996.
XX
PF      28-MAR-1996; 96WO-US01689.
XX
PR      31-MAR-1995; 95US-0414417.
XX
PA      (UNIW ) UNIV WASHINGTON.
XX
PI      Cheever MA, Disis ML;
XX
WPI: 1996-455361/45.
P-PSDB: AAW01111.
XX
DNA encoding HER-2-neu poly:peptide(s) - used for prevention or
treatment of malignancies with which the HER-2/neu oncogene is
associated
XX
Claim 1; Page 49-56; 7lpp; English.
XX
Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or
c-erbB2) protein (AAW01111). The oncogene is overexpressed in various
cancers, including breast, ovarian, colon, lung and prostate, and
appears to induce malignancies through quantitative mechanisms that
result from increased or deregulated expression of an essentially
normal gene product. Nucleotides 2026-3765 of the cDNA sequence
code for the intracellular domain (Iys676-Vall255) of the HER-2/neu
protein, which is useful for immunisation against malignancy.
XX
Nucleic acids can be used to direct expression of the intracellular
domain in transformed host cells, or are used, alone or in a viral
vector, for genetic immunisation of an animal.
XX
SQ      Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 5 97e-32 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

```

Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
SEQ5 (1-59) x AAT40739 (1-3768)
QY      1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
      |||||
Db      2968 CAGAATGAGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTACCGCTCAGTCTG 3027
QY      21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
      |||||
Db      3028 GAGGACGATGACATGGGGACCTGGTGGATCTGAGAGATATCTGTGATCCCGACGAGGC 3087
QY      41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
      |||||
Db      3088 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGTGTCCACACAGGCACCGC 3144

RESULT 9
AAX01912
ID      AAX01912 standard; DNA; 3768 BP.
XX
AC      AAX01912;
XX
DT      21-APR-1999 (first entry)
XX
DE      Human HER-2/neu oncogene DNA.
XX
KW      HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
KW      malignancy; treatment; tumour; ss.
XX
OS      Homo sapiens.
FH      Key
FT      CDS
      1..3768
      /*tag= a
      /product= "HER-2/neu"
      /note= "Oncogene"
      2026..3765
      /*tag= b
      /note= "region which elicits immune response"
XX
PN      US5869445-A.
XX
PD      09-FEB-1999.
XX
PF      01-APR-1996; 96US-0625101.
XX
PR      01-APR-1996; 96US-0625101.
PR      17-MAR-1993; 93US-0033644.
PR      12-AUG-1993; 93US-0106112.
PR      31-MAR-1995; 95US-0414417.
XX
PA      (UNIW ) UNIV WASHINGTON.
XX
PI      Cheever MA, Disis ML;
XX
WPI: 1999-152835/13.
P-PSDB: AAW92406.
XX
Use of HER-2/neu polypeptides - for eliciting an immune response to
an HER-2/neu associated malignancy, particularly for treating or
preventing tumours
XX
Claim 1a; Column 23-32; 26pp; English.
XX
This sequence encodes the human HER-2/neu oncogene protein. A fragment
of this protein is used in a method for eliciting or enhancing an immune
response to HER-2/neu protein. The polypeptide can stimulate T cells and
B cells to produce an immune response to the HER-2/neu protein. The
method can be used for immunisation against a malignancy in which the
HER-2/neu oncogene is associated and in the treatment of an existing
tumour, or to prevent tumour occurrence or reoccurrence.
XX

```


SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 5,97e-32 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

SEQ5 (1-59) x AAX01912 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGCGGC 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGCTGCCACACAGCAGCGGC 3144

RESULT 10

AAA09455
ID AAA09455 standard; DNA; 3768 BP.

XX AC AAA09455;

XX DT 10-AUG-2000 (first entry)

XX DE Human heregulin 2 (Her2) coding sequence.

XX KW Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
KW self-protein; cell-associated peptide antigen; foreign epitope;
KW cancer; breast cancer; prostate cancer; ss.

XX OS Homo sapiens.

XX PN WO200020027-A2.

XX PD 13-APR-2000.

XX PF 05-OCT-1999; 99WO-DK00525.

XX PR 05-OCT-1998; 98DK-0001261.

XX PR 20-OCT-1998; 98US-0105011.

XX PA (MEBI-) M & E BIOTECH AS.

XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;

XX WPI: 2000-349917/30.

XX DR P-PSDB; AAY92620.

XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer

XX Claim 62; Page 187-193; 220pp; English.

XX The claims detail a method for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, human
CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and

CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively.

XX SQ Sequence 3768 BP; 759 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 5,97e-32 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

SEQ5 (1-59) x AAA09455 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGCGGC 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGCTGCCACACAGCAGCGGC 3144

RESULT 11

AAH23392

ID AAH23392 standard; DNA; 3768 BP.

XX AC AAH23392;

XX DT 25-SEP-2001 (first entry)

XX DE Human HER-2/neu protein encoding DNA.

XX KW Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 1..3768

XX FT /*tag= a

XX FT /product= "HER-2/neu protein"

XX PN WO200153463-A2.

XX PD 26-JUL-2001.

XX PF 19-JAN-2001; 2001WO-US01850.

XX PR 21-JAN-2000; 2000US-0177545.

XX PA (CORI-) CORIXA CORP.

XX PI Cheever MA, Hand-Zimmermann S;

XX WPI: 2001-476112/51.

XX DR P-PSDB; AAB85458.

XX New antigen-presenting cells, useful as vaccines for eliciting or
PT enhancing an immune response to HER-2/neu protein, particularly useful
PT for treating or preventing cancer, e.g. breast cancer -

PS Claim 1; Page 41-46; 49pp; English.

XX The invention provides an isolated antigen-presenting cell, which
CC expresses at least an immunogenic portion of a polypeptide that produces
CC an immune response to HER-2/neu protein. The antigen-presenting cells are

CC useful as vaccines for eliciting or enhancing an immune response to
 CC HER-2/neu protein, particularly in treating or preventing malignancies in
 CC which the HER-2/neu oncogene is associated. Specifically, these are
 CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,
 CC colon, lung or prostate cancers. The present sequence represents a DNA
 CC encoding the human HER-2/neu protein (also known as p185 or c-erbB2).
 XX

SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
 Pred. No.: 5.97e-32 Length: 3768
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

SEQ5 (1-59) x AAH23392 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 |||||
 DB 2968 CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACGCTG 3027
 |||||
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
 |||||
 DB 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTATACCCAGCAGGC 3087
 |||||
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
 |||||
 DB 3088 TTCTTCTGTCCAGACCTGCCCCGGGGCGTGGGGGCGATGTCACACAGGCACCGC 3144

RESULT 12

ABZ35744

ID ABZ35744 standard; DNA; 3768 BP.

XX AC ABZ35744;

XX DT 07-FEB-2003 (first entry)

XX DE Human ERBB2 polynucleotide SEQ ID NO 52.

XX DE Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
 KW protozoacide; gene expression; antisense; tumour; infection; plasmodium;
 KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
 KW Hepatitis C virus; human papilloma virus; gene; ds.

OS Homo sapiens.

XX PN DE10100588-Al.

XX PD 18-JUL-2002.

XX PF 09-JAN-2001; 2001DE-1000588.

XX PR 09-JAN-2001; 2001DE-1000588.

XX PA (RIBO-) RIBOPHARMA AG.

XX PI Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX DR WPI; 2002-683450/74.

XX PT Inhibiting expression of target genes, useful e.g. for treating tumors,
 PT by introducing into cells two double-stranded RNAs that are
 PT complementary to the target
 XX Claim 13; Page 38-39; 100pp; German.

XX The invention relates to inhibiting expression of a target gene in a cell
 CC by introducing at least two oligoribonucleotides (dsRNAi and II), both
 CC with a double-stranded (ds) structure of at most 49 sequential nucleotide
 CC pairs. At least part of one strand (S1, S2) of the ds structures in each
 CC of dsRNAi and II are complementary to regions in the target gene. The

CC method uses antisense inhibition of gene expression using double stranded
 CC RNA inhibition (RNAi). The method is particularly used to treat tumours
 CC or infections, especially by plasmodium or viruses/viroids (pathogenic on
 CC humans, animals or plants). The method provides more effective inhibition
 CC of expression than known methods using a single dsRNA, even at very low
 CC concentrations. When dsRNA has at least one unpaired nucleotide at the
 CC end, stability (and thus effective concentration in the cell) is
 CC improved and efficiency can be increased further by pretreating the cells
 CC with interferon. The present sequence is that of a target DNA of the
 CC invention.

XX SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
 Pred. No.: 5.97e-32 Length: 3768
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

SEQ5 (1-59) x ABZ35744 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 |||||
 DB 2968 CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACGCTG 3027
 |||||
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
 |||||
 DB 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTATACCCAGCAGGC 3087
 |||||
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
 |||||
 DB 3088 TTCTTCTGTCCAGACCTGCCCCGGGGCGTGGGGGCGATGTCACACAGGCACCGC 3144

RESULT 13

ABX09987

ID ABX09987 standard; DNA; 3768 BP.

XX AC ABX09987;

XX DT 23-JAN-2003 (first entry)

XX DE Human ERBB2 DNA fragment SEQ ID 52.

XX DE Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;
 KW prion; inhibition; human; ds.

XX OS Homo sapiens.

XX PN DE10100587-cl.

XX PD 21-NOV-2002.

XX PF 09-JAN-2001; 2001DE-1000587.

XX PR 09-JAN-2001; 2001DE-1000587.

XX PA (RIBO-) RIBOPHARMA AG.

XX PI Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX DR WPI; 2002-742209/81.

XX PT Inhibiting expression of target genes, e.g. oncogenes, in cells, by
 PT introduction of complementary double-stranded oligoribonucleotide,
 PT after treating the cell with interferon
 XX Disclosure; Page 43-44; 98pp; German.

XX This invention describes a novel method for inhibiting expression of a
 CC target gene by introducing into the cell that contains the target gene
 CC at least one oligoribonucleotide (dsRNAi) that has a double-stranded

CC (ds) structure of not more than 49 consecutive nucleotides (nt), where
 CC at least a segment of one strand of the ds structure is complementary
 CC with the target gene and the cells are treated with interferon before
 CC introduction of dsRNA1. The method is used to inhibit expression of
 CC target genes, particularly oncogenes, cytokine genes, Id (not defined)
 CC protein genes; developmental or prion genes, or genes expressed in
 CC pathogenic organisms (particularly plasmodia) or in viruses or viroids
 CC (pathogenic in humans, animals or plants). Treating the cells with
 CC interferon greatly increases the extent to which dsRNA can inhibit
 CC expression of the target genes, and the effect is even greater when dsRNA
 CC are modified to increase their stability. ABX09936-ABX10075 represent
 CC gene fragments used to illustrate the method of the invention.

XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
 Pred. No.: 5,97e-32 Length: 3768
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

SEQ5 (1-59) x ABX09987 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTACTGCTG 3027

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
 Db 3028 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3087

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
 Db 3088 TTCCTCTCCAGACCTGCCCCGGGCGCTGGGGCGATGGTCCACACAGCACCGC 3144

RESULT 14

AAAD43935
 ID AAD43935 standard; cDNA; 3768 BP.

XX AAD43935;

DT 13-DEC-2002 (first entry)

DE Human HER-2 cDNA.

XX Transgenic animal; transgenic; mammary gland cell; HER2; tumour;
 KW cancer; therapy; apoptosis; cytostatic; human; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..3768
 FT /*tag= a
 FT /product= "Human HER2 protein"

XX US200203736-A1.

XX 21-MAR-2002.

XX 16-MAR-2001; 2001US-0811115.

XX 16-MAR-2000; 2000US-189844P.

XX (ERIC/) ERICKSON S.

PA (KING/) KING K.

PA (SCHW/) SCHWALL R.

XX Erickson S, King K, Schwall R;

XX WPI; 2002-401155/43.

DR P-PSDB; AAE26349.

XX New transgenic non-human mammal that produces detectable levels of a
 PT native human HER2 protein in its mammary gland cells, useful as tumor
 PT models for testing HER2-directed cancer therapies, and for identifying
 PT anticancer agents

PS Example 2; Page 24-26; 83pp; English.

XX The invention relates to a transgenic non-human mammal that produces in
 CC its mammary gland cells detectable levels of a native human HER2 protein
 CC or its fragment. The transgenic animals are useful as tumour models for
 CC testing HER2-directed cancer therapies, and for identifying anticancer
 CC agents. The animals may also be used as source of cells which can be
 CC immortalised in culture, in screening for compounds that have potential
 CC as prophylactic or therapeutic treatments of diseases or disorders
 CC involving expression of HER2. The anti-cancer molecules are useful for
 CC inducing apoptosis or cell death of cancer cells. The present sequence
 CC is human HER-2 cDNA.

SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
 Pred. No.: 5,97e-32 Length: 3768
 Score: 326.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

SEQ5 (1-59) x AAD43935 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTACTGCTG 3027

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
 Db 3028 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3087

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
 Db 3088 TTCCTCTCCAGACCTGCCCCGGGCGCTGGGGCGATGGTCCACACAGCACCGC 3144

RESULT 15

AAAD43986
 ID AAD43986 standard; DNA; 3768 BP.

XX AAD43986;

DT 13-DEC-2002 (first entry)

DE Human Her2 antigen DNA.

XX Human; immune response; T-helper cell epitope; chitosan; CTL response;
 KW vaccine; prostate cancer; breast cancer; Her2 antigen; cytostatic;
 KW immunostimulant; gene; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..3768
 FT /*tag= a
 FT /product= "Human Her2 antigen"
 FT sig_peptide 1..69
 FT /*tag= b
 FT mat_peptide 70..2250
 FT /*tag= c
 FT /product= "Mature human Her2 antigen"

XX WO200234287-A2.

XX 02-MAY-2002.

XX

```

PF 26-OCT-2001; 2001WO-DK00705.
XX
PR 27-OCT-2000; 2000DK-0001606.
PR 03-NOV-2000; 2000US-245166P.
PR 18-JUN-2001; 2001DK-0000936.
XX
PA (PHAR-) PHARMEXA AS.
XX
XX Beier AM, Gautam A, Mouritsen S;
PI
XX
XX WPI; 2002-463339/49.
DR P-PSDB; AAE26366.
XX
XX Inducing or enhancing an immune response against an antigen,
PT particularly cytotoxic T-lymphocyte responses, for treating or
PT ameliorating prostate or breast cancer, comprises administering the
PT antigen formulated with chitosan -
XX
XX Claim 28; Page 85-90; 97pp; English.
PS
XX The invention relates to a method for inducing or enhancing an immune
CC response against a polypeptide antigen in an animal, including an immune
CC The method comprises administering the polypeptide antigen or at least
CC one variant which includes at least one first T-helper cell epitope that
CC is foreign to the animal (foreign TH epitope) and is formulated with
CC chitosan. The polypeptide antigen is weakly immunogenic or non-
CC immunogenic. The invention is used as vaccine. The chitosan and
CC polypeptide antigen or its variant are useful in the preparation of an
CC immunogenic composition for inducing or enhancing an immune response,
CC particularly CTL response, against the polypeptide or protein antigen.
CC The method for inducing or enhancing an immune response is useful in
CC treating or ameliorating cancer, e.g. prostate or breast cancer. The
CC present sequence is human Her2 antigen DNA.
XX
XX SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 5.97e-32 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ5 (1-59) x AAD43986 (1-3768)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGGACITGGGCCCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGGTG 3027
QY 21 GluAspAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATCAGATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACCACAGGCACCCG 3144

```

Search completed: October 15, 2003, 23:50:11
Job time : 61.0384 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 849.651 Seconds
(without alignments)
2840.777 Million cell updates/sec

Title: SEQ5
Perfect score: 326
Sequence: 1 QNEDLGASPLDSTFYRSL.....GFFCPDPAPGAGMVHRRH 59

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool_p/HOLLERAN480/runat_15102003_131912_20536/app_query.fasta_1.4685
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480.qcgn.1.1.2724.@runat_15102003_131912_20536 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	326	100.0	1755	6	AX384609 Sequence
2	326	100.0	1767	6	AX384607 Sequence
3	326	100.0	1773	6	AX384610 Sequence
4	326	100.0	1806	6	AX384608 Sequence
5	326	100.0	3678	6	AX505114 Sequence
6	326	100.0	3768	6	AR034479 Sequence
7	326	100.0	3768	6	AX060704 Sequence
8	326	100.0	3768	6	AX201817 Sequence
9	326	100.0	3768	6	AX380923 Sequence
10	326	100.0	3768	6	AX384604 Sequence
11	326	100.0	3768	6	AX465456 Sequence
12	326	100.0	3768	6	AX467229 Sequence
13	326	100.0	3768	6	AX481438 Sequence
14	326	100.0	4473	6	AR080259 Sequence
15	326	100.0	4473	6	AR167390 Sequence
16	326	100.0	4473	9	HSERB2R
17	326	100.0	4530	6	AR202597 Sequence
18	326	100.0	4530	6	AR283481 Sequence
19	326	100.0	4530	6	AX282577 Sequence
20	326	100.0	4530	6	AX587649 Sequence
21	326	100.0	4530	6	AX644071 Sequence
22	326	100.0	4530	6	BD005474 Cellular
23	326	100.0	4530	6	I21124 Sequence 9
24	326	100.0	4530	6	I59745 Sequence 9
25	326	100.0	4530	6	HMHER2A
26	326	100.0	9274	6	AX060703 Sequence
27	326	100.0	13450	9	AB096612 Homo sapi
28	326	100.0	16572	9	AB096614 Homo sapi
29	326	100.0	20271	9	AB096613 Homo sapi
30	326	100.0	30837	9	AY208911 Homo sapi
31	326	100.0	161815	9	AC079199 Homo sapi
32	326	100.0	168585	9	AC040933 Homo sapi
33	326	100.0	198008	2	AC142197 Homo sapi
34	300	92.0	3780	4	AB008451 Canis fam
35	288	88.3	4062	10	D16295 Mesocricetu
36	279	85.6	3955	6	AX380924 Sequence
37	279	85.6	3955	6	I21129 Sequence 14
38	279	85.6	3955	6	I59750 Sequence 14
39	279	85.6	3955	10	X03362 Rat mRNA fo
40	279	85.6	4727	10	AY116182 Rattus no
41	274	84.0	23732	2	AC098491 Rattus no
42	269	82.5	2083	10	BC027080 Mus muscu
43	269	82.5	2763	6	AX380942 Sequence
44	269	82.5	2781	6	AX380944 Sequence
45	269	82.5	3771	6	AX189662 Sequence

ALIGNMENTS

```

AX384609
LOCUS AX384609 1755 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 6 from Patent WO0214503.
ACCESSION AX384609
VERSION AX384609.1 GI:19577810
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 6 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..1755
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 376 a 517 c 531 g 331 t
ORIGIN

Alignment Scores:
Pred. No.: 1.9e-31 Length: 1755
Score: 326.00 Matches: 59
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0

SEQ5 (1-59) x AX384609 (1-1755)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 946 CAGAATGAGGACTTGGCCCGAGCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 1006 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGTGTACCCAGCAGGC 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
Db 1066 TTCTTCTGTCAGACCTGCCCCGGGGCTGGGGGCGATGTTCCACACAGCACCGC 1122

RESULT 2
AX384607
LOCUS AX384607 1767 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 4 from Patent WO0214503.
ACCESSION AX384607
VERSION AX384607.1 GI:19577808
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 4 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..1767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 381 a 521 c 529 g 336 t
ORIGIN

AX384607
LOCUS AX384607 1773 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 7 from Patent WO0214503.
ACCESSION AX384610
VERSION AX384610.1 GI:19577811
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 7 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..1773
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 383 a 528 c 530 g 332 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1.92e-31 Length: 1767
Score: 326.00 Matches: 59
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0

SEQ5 (1-59) x AX384607 (1-1767)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 946 CAGAATGAGGACTTGGCCCGAGCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 1006 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGTGTACCCAGCAGGC 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
Db 1066 TTCTTCTGTCAGACCTGCCCCGGGGCTGGGGGCGATGTTCCACACAGCACCGC 1122

RESULT 3
AX384610
LOCUS AX384610 1773 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 7 from Patent WO0214503.
ACCESSION AX384610
VERSION AX384610.1 GI:19577811
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 7 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..1773
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 383 a 528 c 530 g 332 t
ORIGIN

Alignment Scores:
Pred. No.: 1.92e-31 Length: 1773
Score: 326.00 Matches: 59
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0

SEQ5 (1-59) x AX384610 (1-1773)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 970 CAGAATGAGGACTTGGCCCGAGCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGCTG 1029
QY 21 GluAspAspMetGlyAspLeuValAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 1030 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGTGTACCCAGCAGGC 1089
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
Db 1090 TTCTTCTGTCAGACCTGCCCCGGGGCTGGGGGCGATGTTCCACACAGCACCGC 1146

RESULT 4
AX384608
LOCUS AX384608 1806 bp DNA linear PAT 19-MAR-2002

```

```

1. 3678 /organism="Homo sapiens"
/ncbi_type="genomic DNA"
/db_xref="taxon:9606"
7. .3678
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD48804.1"
/db_xref="GI:23886422"

```



```

DB: 6 Gaps: 0

SEQ5 (1-59) x AR034479 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
    |||||
Db 2968 CAGAAAGAGGACTTGGGGCCAGCCAGTCCCTTGGACAGCAGCACCTTCTACCGCTCAGCTGCTG 3027
    |||||
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40
    |||||
Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCACGAGGCG 3087
    |||||
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
    |||||
Db 3088 TTCTTCTGTCCAGACCTGCGCCGGCGCTGGGGGCGATGCTCCACACGAGGACCGC 3144

RESULT 7
AX060704
LOCUS AX060704 3768 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 2 from Patent WO0100244.
ACCESSION AX060704
VERSION AX060704.1 GI:12406101
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Erickson, S. and Schwall, R.
TITLE Methods of treatment using anti-erbB antibody-maytansinoid
            conjugates
JOURNAL Patent: WO 0100244-A 2 04-JAN-2001;
            Genentech, Inc. (US)
FEATURES
    source
        1. .3768
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 4.41e-31 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

SEQ5 (1-59) x AX060704 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
    |||||
Db 2968 CAGAAAGAGGACTTGGGGCCAGCCAGTCCCTTGGACAGCAGCACCTTCTACCGCTCAGCTGCTG 3027
    |||||
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40
    |||||
Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCACGAGGCG 3087
    |||||
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
    |||||
Db 3088 TTCTTCTGTCCAGACCTGCGCCGGCGCTGGGGGCGATGCTCCACACGAGGACCGC 3144

RESULT 8
AX201817
LOCUS AX201817 3768 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0153463.
ACCESSION AX201817
VERSION AX201817.1 GI:15391666
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Cheever, M.A. and Hand-Zimmermann, S.
Compounds and methods for prevention and treatment of her-2/ neu
associated malignancies
Patent: WO 0153463-A 1 26-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
    Location/Qualifiers
        source
            1. .3768
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
            1. .3768
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAC60126.1"
                /db_xref="GI:15391667"
                /translation="MELAAALRWGLLLALLPPGAASQTQVCTGTDMMKLRLLPASPETHLD
                MLRHLYQGVGVGNLELYLPTNASLSFLOIEVOGYVLIHAHQVQVPLQRLRIV
                RGTQLFDNYALVLDNGDPLNNTTPTVGASPGGLRELQRLSLTELKGLVLIQRNPQ
                LCYQDTILWKDIFHKNOLALTLIDTNRSRACHPCSPMCKSGRCESEDCQSLTRT
                VCAGGACRCKGPLPTDCCHEOCAAGCTGPKHSDCLACILHNHSGICELACPALVTYT
                DIFESMPNPEGRTFGASCVTACPYNLSTDVGSCTLVCLHNQEVTAEDGTRCEKC
                SKPCARVCYGLGMEHLREVRVTSANIQEFAGCKIFGLSLAFEPESFDGDPASNTAPL
                QPEOLFETLEEITGYLYISAMPDLSLDLVFNQQLVIRGLHNHNGAYSLTLQGLGI
                SWLGRSLRELGLALIHNNTHLCFVHTVPWDQLFNPHQALLTANRPEDECVEG
                LACHOLCARGHCWGPGTQCVNCSQFLRGQECVECRVLQGLPREVYNARHCLPCHE
                CPGNGSVTCFGEADQCVACAHYKDPFPCVARGSVKPDLSYMPIMKFPDEEGACQ
                PGFINCHSDVLDKDGKPAEQRASPLTSISANVVGILLVVLGVVGLILIKRROKI
                RYKTMRLIQETELVEPLTPSGAMPNQAMRILKTELKRVXVLGASGAGTGYKGIWI
                PDGENVKIPVAIKVRENTSPRANKEILDEAYVMAGVSPYSVRLIGLICLTSTVQLVT
                QLMYGCLLDHVRENRLGSQLNMCMIAYKMSYLEDVRLVHRLDAARNVLYKSP
                NHVKITDFGLARLLDIDETEHADGKVPKWMALESILRR3FTHQSDVMXGYVTWE
                LMTFGAPYDGIAPAREIPDLLEKGERLPDPTCTIDVYIMYKWMIDSECRPRREL
                VSEFSRMAPDQRFVYVIONEDLGASPLDSTFYSLLEDDMDGLVDAEYLVPOQGF
                FCPDPAPGAGGAWHHRHSSTRSGGDLTLGLEPSEEARPSPLASGASGSDVDG
                DLGMAAGLQSLDHPSPLOYSIEDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVR
                PQPPSPREGPLFAARPAATLRLPKTLSPKNGVKVDYFAFGGAVENPEYLTLPQGAA
                POPHPPPAFSPAFDNLVYWDQDPPPERGAPPSTFKGTPTAENPEYGLDVPV"
BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 4.41e-31 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

SEQ5 (1-59) x AX201817 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
    |||||
Db 2968 CAGAAAGAGGACTTGGGGCCAGCCAGTCCCTTGGACAGCAGCACCTTCTACCGCTCAGCTGCTG 3027
    |||||
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40
    |||||
Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCACGAGGCG 3087
    |||||
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
    |||||
Db 3088 TTCTTCTGTCCAGACCTGCGCCGGCGCTGGGGGCGATGCTCCACACGAGGACCGC 3144

RESULT 9
AX380923
LOCUS AX380923 3768 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 9 from Patent WO0212341.
ACCESSION AX380923
VERSION AX380923.1 GI:19575767
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Cheever, M.A. and Gheysen, D.
Her-2/neu fusion proteins
Patent: WO 0212341-A 9 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
Location/Qualifiers
1. .3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1. .3768
/note="human Her-2/neu protein"
/codon_start=1
/protein_id="CAD28392.1"
/db_xref="GI:19575768"
/translation="MELAAICRWGLLLALLPPGAASQVCTGTDMLRLPASPETHLD
MLRLHYGCVQVQGNLELYLPTNASLSFLQDIOEVQGYVLIHNRQVPLQLRLIV
RGTLQFEDNYALVDNGDPLNNTPTVTGASPGGLRELQLRSLTEILKGGVLIQRNPQ
LCYQDTILWKDIFHKHNLALTLIDTNRSRACHPCSPMCKGRCWGESSEDCQSLRT
VCAGCARCKGKPLTDCHEQCAAGCTGPKHSDCLACLHPNHSGICELHCPALVYNT
DTFESMPNPEGRTYFGASCYTACPNYLSVDGCTLVCLPHNQEVTAEDGTQCEKC
SKPCARVCYIGLMEHLREVRVTSANTIOEFAGCKIFGSLAFPESEFDGDPASNTAPL
QPEQLQVFETLEEITGYLISAWPDSLPLDSVFQNLQVIRGRILHNGAYSLSLQGLGI
SWLGLRSLRELGSGLALIHNTHLCFVHTVPMDFQNLFRPHQALLHTANRPEDECVGG
LACHOLCARHCWGPGTQCVCNCSQFLRGQCEVCEVRVLOGLPREYVNRHCLPCHPE
COPONGVTCFGEADOCVACAHYKDPFCVACPCSPGVKPDLSYMPWKPPDEGACQ
PCPINCHSCVDLDDKCPAQRASPLTISIISAVGILLVVLGVVGLIKRQOKI
RKYTMRLRLQETELVEPFGAMPNQAKMILKELTKRVKVLGSGAFGVYKGIWI
PDGENVIPAIVKIRENTSPKANKEILDVAYVMAGVSPVYSLRLGICLTSTVLQVLT
OLMPYGLLDHVRNRLGASQDLNNCMQIAGMSYLEDVRLVHRDLAARNVLYKSP
NHVITDFGLARLDIDETEVHADGGKVPKIMWALESILRRRFTHQSDVNSYGVYVWE
LMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMLVWKCMWIDSECRPFREL
VSESRWARDPQRVVIQNEGLGPASPLDSTFYSLLEDMDGLVDAEYLVYQQGF
FCDDPAAGAGMWHHRSSSTRSGGDLTLGLEPSEEEAPRSLAPSEAGSGSVDFG
DLGMAAGLQSLTPHDPSPQRYSEDTPLPSETDGYVAPLTCSPQEVYNQDVR
POPPSPREGPLPAAPAGATLERPKTLSPGKNGVYKDVAFGGAVERNPEYLTPOGGAA
PQPHPPAFSPAFDNLYYWDQDPPERGAPSTFKGTPTAENPEYLGLDVPV"
1. .1959
/note="extracellular domain (ECD) of human Her-2/neu"
2026. .3765
/note="extracellular domain (ICD) of human Her-2/neu"
2968. .3765
/note="phosphorylation domain (PD) of human Her-2/neu"
2968. .3144
/note="preferred portion of the phosphorylation domain
(delta PD) of human Her-2/neu"
BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 4.41e-31 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ5 (1-59) x AX380923 (1-3768)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCCTCTGTCAGACCTTGGCCCGGGGCGTGGGGGCGATGGTCCACCAGCAGCACCGC 3144

RESULT 10
AX384604
LOCUS AX384604 3768 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 1 from Patent WO0214503.
ACCESSION AX384604
VERSION AX384604.1 GI:19577806
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D., McNeill, P.D. and Vedvick, T.S.
Compositions and methods for the therapy and diagnosis of her-2/neu-associated malignancies
Patent: WO 0214503-A 1 21-FEB-2002;
CORIXA CORPORATION (US)
Location/Qualifiers
1. .3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1. .3768
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD28577.1"
/db_xref="GI:19577807"
/translation="MELAAICRWGLLLALLPPGAASQVCTGTDMLRLPASPETHLD
MLRLHYGCVQVQGNLELYLPTNASLSFLQDIOEVQGYVLIHNRQVPLQLRLIV
RGTLQFEDNYALVDNGDPLNNTPTVTGASPGGLRELQLRSLTEILKGGVLIQRNPQ
LCYQDTILWKDIFHKHNLALTLIDTNRSRACHPCSPMCKGRCWGESSEDCQSLRT
VCAGCARCKGKPLTDCHEQCAAGCTGPKHSDCLACLHPNHSGICELHCPALVYNT
DTFESMPNPEGRTYFGASCYTACPNYLSVDGCTLVCLPHNQEVTAEDGTQCEKC
SKPCARVCYIGLMEHLREVRVTSANTIOEFAGCKIFGSLAFPESEFDGDPASNTAPL
QPEQLQVFETLEEITGYLISAWPDSLPLDSVFQNLQVIRGRILHNGAYSLSLQGLGI
SWLGLRSLRELGSGLALIHNTHLCFVHTVPMDFQNLFRPHQALLHTANRPEDECVGG
LACHOLCARHCWGPGTQCVCNCSQFLRGQCEVCEVRVLOGLPREYVNRHCLPCHPE
COPONGVTCFGEADOCVACAHYKDPFCVACPCSPGVKPDLSYMPWKPPDEGACQ
PCPINCHSCVDLDDKCPAQRASPLTISIISAVGILLVVLGVVGLIKRQOKI
RKYTMRLRLQETELVEPFGAMPNQAKMILKELTKRVKVLGSGAFGVYKGIWI
PDGENVIPAIVKIRENTSPKANKEILDVAYVMAGVSPVYSLRLGICLTSTVLQVLT
OLMPYGLLDHVRNRLGASQDLNNCMQIAGMSYLEDVRLVHRDLAARNVLYKSP
NHVITDFGLARLDIDETEVHADGGKVPKIMWALESILRRRFTHQSDVNSYGVYVWE
LMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMLVWKCMWIDSECRPFREL
VSESRWARDPQRVVIQNEGLGPASPLDSTFYSLLEDMDGLVDAEYLVYQQGF
FCDDPAAGAGMWHHRSSSTRSGGDLTLGLEPSEEEAPRSLAPSEAGSGSVDFG
DLGMAAGLQSLTPHDPSPQRYSEDTPLPSETDGYVAPLTCSPQEVYNQDVR
POPPSPREGPLPAAPAGATLERPKTLSPGKNGVYKDVAFGGAVERNPEYLTPOGGAA
PQPHPPAFSPAFDNLYYWDQDPPERGAPSTFKGTPTAENPEYLGLDVPV"
BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 4.41e-31 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ5 (1-59) x AX384604 (1-3768)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCCTCTGTCAGACCTTGGCCCGGGGCGTGGGGGCGATGGTCCACCAGCAGCACCGC 3144

```

Db      3088 TTCTTCTGTCCAGACCCCTGCCCGGGCGCTGGGGGCATGTCACACAGGCACCGC 3144
RESULT 11
AX465456
LOCUS   AX465456                      3768 bp    DNA        linear    PAT 16-JUL-2002
DEFINITION
Sequence 1 from Patent WO0213847.
ACCESSION
AX465456
VERSION
AX465456.1 GI:21899819
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Gaiger, A., Cheever, M.A. and Hand-Zimmermann, S.
TITLE
Methods for diagnosis and therapy of hematological and
virus-associated malignancies
JOURNAL
Patent: WO 0213847-A 1 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
Location/Qualifiers
source
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD42363.1"
/db_xref="GI:21899820"
/translation="MELAAALRWGLLLALLPPGAASQTVCVTGDMKRLRPLASPETHLD
MLRHLYQGVQVGNLELTYPNTNASLFDQIEQVQGYVLIHNNQVQVPLQRLRIV
RGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGRLRLQRLSLTELKGGVLIQRNPQ
LCVQDTILWKDILFHNKQALTLIDTNRSRACHPCSPCKGSRGSESSDQSLTFT
VCAGGACGKGLPTDCCHEQCAAGCTGPKHSDCLACLHNSGICELHCPALVTYT
DTFESMPNPEGRTFGASCVTACPYNLSTDVSGCTVCLPLNQEVTDGQRCCK
SPKCARCYGLGMEHLREVRATVSANIQEFAGCKIFGSLAFLPESFDGDPASNTAPL
QPQLQVFTLEETGYLISAMPDSDLPSVFNQLQVIRGRILHNGAYSLTIQGLGI
SWGLRSRLGSLGALIHNTHLCFVHTVPMDQLFRNPQALLHTANRPEDECVGEG
LACHOLCARGHCWPGPTQCVNCSQFLRGQECVECRVLOGLPREVYNARHCLPCPHE
COPQNSVTCFGEADQCVACAHYKDPFCVACRCPGSKVPDLISYMPWKFPDEEGACO
PCPINTHSCVDLDDKGCFAEQRASPLTISIISAVGILLVVLGVVFGILLIKRQOKI
RYTMRRLQELTELVELPTPSGAMPNQAMRIKTELKRVKVLGSGAGFTVYKGIWI
PDGENVKIPVAIKLRENTSPKANKEILDEAYVMAGVSPYVSRLLGICLTSTVQLVT
OLMPYGLLDHVRNENRGLSQDLNMCQIAKMSYLEDVRLVHDLAARNVLYKSP
NHVKTIDFGLARLIDDEYTHADGKVPDKWMALESILRRRFTHOSDVMXGVTWE
LMTFGAKPYDGIIPAREIPDLLEKGERLPPOPICTIDVYIMVKWCWMDSECRPREL
VSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDMDGLVDACEYLPQGF
FCPDPAPAGGWMVHHRSSSTRSGGDLTLGLEPSEEAAPRSLAPSGAGSDVPDG
DLGMAAGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVR
POPSPREGPLPAARPAAGATLERPKTLSPKNGVVKDVFAGGAVENPEYLTPOGGAA
POPHPPPAFSPAFDNLXYWDQDPPPERGAPSTFKGTPTAENPEYLGLDVVP"

BASE COUNT      759 a 1171 c 1119 g 719 t
ORIGIN

Alignment Scores:
Pred. No.:      4,41e-31      Length:      3768
Score:          326.00      Matches:      59
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              6              Gaps:      0

SEQ5 (1-59) x AX465456 (1-3768)

QY      1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
|||||
Db      2968 CAGAAATGAGACTTGGGCCAGCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
|||||
QY      21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
|||||
Db      3028 GAGGACGATGACATGGGGACCTGGTGATGCTGAGAGTATCTGGTACCCACGACGCGC 3087
|||||
QY      41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59

```

```

Db      3088 TTCTTCTGTCCAGACCCCTGCCCGGGCGCTGGGGGCATGTCACACAGGCACCGC 3144
RESULT 12
LOCUS   AX467229                      3768 bp    DNA        linear    PAT 16-JUL-2002
DEFINITION
Sequence 3 from Patent WO0234287.
ACCESSION
AX467229
VERSION
AX467229.1 GI:21900511
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Beier, A.M., Gautam, A. and Mouritsen, S.R.
TITLE
Novel therapeutic vaccine formulations
JOURNAL
Patent: WO 0234287-A 3 02-MAY-2002;
Pharmexa A/S (DK)
FEATURES
Location/Qualifiers
source
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD42452.1"
/db_xref="GI:21900512"
/translation="MELAAALRWGLLLALLPPGAASQTVCVTGDMKRLRPLASPETHLD
MLRHLYQGVQVGNLELTYPNTNASLFDQIEQVQGYVLIHNNQVQVPLQRLRIV
RGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGRLRLQRLSLTELKGGVLIQRNPQ
LCVQDTILWKDILFHNKQALTLIDTNRSRACHPCSPCKGSRGSESSDQSLTFT
VCAGGACGKGLPTDCCHEQCAAGCTGPKHSDCLACLHNSGICELHCPALVTYT
DTFESMPNPEGRTFGASCVTACPYNLSTDVSGCTVCLPLNQEVTDGQRCCK
SPKCARCYGLGMEHLREVRATVSANIQEFAGCKIFGSLAFLPESFDGDPASNTAPL
QPQLQVFTLEETGYLISAMPDSDLPSVFNQLQVIRGRILHNGAYSLTIQGLGI
SWGLRSRLGSLGALIHNTHLCFVHTVPMDQLFRNPQALLHTANRPEDECVGEG
LACHOLCARGHCWPGPTQCVNCSQFLRGQECVECRVLOGLPREVYNARHCLPCPHE
COPQNSVTCFGEADQCVACAHYKDPFCVACRCPGSKVPDLISYMPWKFPDEEGACO
PCPINTHSCVDLDDKGCFAEQRASPLTISIISAVGILLVVLGVVFGILLIKRQOKI
RYTMRRLQELTELVELPTPSGAMPNQAMRIKTELKRVKVLGSGAGFTVYKGIWI
PDGENVKIPVAIKLRENTSPKANKEILDEAYVMAGVSPYVSRLLGICLTSTVQLVT
OLMPYGLLDHVRNENRGLSQDLNMCQIAKMSYLEDVRLVHDLAARNVLYKSP
NHVKTIDFGLARLIDDEYTHADGKVPDKWMALESILRRRFTHOSDVMXGVTWE
LMTFGAKPYDGIIPAREIPDLLEKGERLPPOPICTIDVYIMVKWCWMDSECRPREL
VSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDMDGLVDACEYLPQGF
FCPDPAPAGGWMVHHRSSSTRSGGDLTLGLEPSEEAAPRSLAPSGAGSDVPDG
DLGMAAGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVR
POPSPREGPLPAARPAAGATLERPKTLSPKNGVVKDVFAGGAVENPEYLTPOGGAA
POPHPPPAFSPAFDNLXYWDQDPPPERGAPSTFKGTPTAENPEYLGLDVVP"

BASE COUNT      758 a 1170 c 1121 g 719 t
ORIGIN

Alignment Scores:
Pred. No.:      4,41e-31      Length:      3768
Score:          326.00      Matches:      59
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              6              Gaps:      0

SEQ5 (1-59) x AX467229 (1-3768)

QY      1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
|||||
Db      2968 CAGAAATGAGACTTGGGCCAGCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
|||||
QY      21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
|||||
Db      3028 GAGGACGATGACATGGGGACCTGGTGATGCTGAGAGTATCTGGTACCCACGACGCGC 3087
|||||
QY      41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59

```

```
Db 3088 TTCTTCTGTCCAGACCTTGCCCGGGCGCTGGGGGCATGGTCCACCAGGCACCGC 3144
|||||
RESULT 13
AX481438
LOCUS AX481438 3768 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 52 from Patent WO20055693.
ACCESSION AX481438
VERSION AX481438.1 GI:22316352
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.
TITLE Method for inhibiting the expression of a target gene
JOURNAL Patent: WO 02055693-A 52 18-JUL-2002;
Ribopharma AG (DE)
FEATURES
source
1. 3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 4,41e-31 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ5 (1-59) x AX481438 (1-3768)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 2968 CAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 3028 GAGGACGATGACATGGGGACCTGGTGGTGTGAGGAGTATCTGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
|||||
Db 3088 TTCTTCTGTCCAGACCTTGCCCGGGCGCTGGGGGCATGGTCCACCAGGCACCGC 3144
|||||
RESULT 14
AR080259
LOCUS AR080259 4473 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 1 from patent US 5968748.
ACCESSION AR080259
VERSION AR080259.1 GI:10006994
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4473)
AUTHORS Bennett, C. Frank., Lipton, A. and Witters, L.M.
TITLE Antisense oligonucleotide modulation of human HER-2 expression
JOURNAL Patent: US 5968748-A 1 19-OCT-1999;
FEATURES
source
1. 4473
/organism="unknown"
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:
Pred. No.: 5,32e-31 Length: 4473
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ5 (1-59) x AR080259 (1-4473)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 3142 CAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 3202 GAGGACGATGACATGGGGACCTGGTGGTGTGAGGAGTATCTGTACCCAGCAGGCG 3261
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
|||||
Db 3262 TTCTTCTGTCCAGACCTTGCCCGGGCGCTGGGGGCATGGTCCACCAGGCACCGC 3318
|||||
Search completed: October 16, 2003, 11:03:09
Job time : 855.833 secs
```

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ5 (1-59) x AR080259 (1-4473)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 3142 CAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 3202 GAGGACGATGACATGGGGACCTGGTGGTGTGAGGAGTATCTGTACCCAGCAGGCG 3261
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
|||||
Db 3262 TTCTTCTGTCCAGACCTTGCCCGGGCGCTGGGGGCATGGTCCACCAGGCACCGC 3318
|||||
RESULT 15
AR167390
LOCUS AR167390 4473 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 26 from patent US 6287569.
ACCESSION AR167390
VERSION AR167390.1 GI:17903168
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4473)
AUTHORS Kipps, T.J. and Wu, Y.
TITLE Vaccines with enhanced intracellular processing
JOURNAL Patent: US 6287569-A 26 11-SEP-2001;
FEATURES
source
1. 4473
/organism="unknown"
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:
Pred. No.: 5,32e-31 Length: 4473
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ5 (1-59) x AR167390 (1-4473)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 3142 CAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 3202 GAGGACGATGACATGGGGACCTGGTGGTGTGAGGAGTATCTGTACCCAGCAGGCG 3261
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
|||||
Db 3262 TTCTTCTGTCCAGACCTTGCCCGGGCGCTGGGGGCATGGTCCACCAGGCACCGC 3318
|||||
Search completed: October 16, 2003, 11:03:09
Job time : 855.833 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 2081.84 Seconds
(without alignments)
3105.423 Million cell updates/sec

Title: SEQ4
Perfect score: 1430
Sequence: 1 QNEDLGPASPLDSTFYRSL.....TFKGTPTAENPEYLGLDVPV 266

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/HOLLERAN480/runat_15102003_131913_20548/app_query.fasta_1.4685
-DB=EST_QFMT-fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=HOLLERAN480 -CGCN_1_12645 @runat_15102003_131913_20548 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -FGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estim:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_fod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1426	98.3	885	14	CA455074	CA455074 AGENCOURT
2	1423	98.1	4715	11	AF318349	AF318349 Homo sapi
3	1399.5	96.5	893	14	CA455141	CA455141 AGENCOURT
4	1354	93.4	894	14	CA454131	CA454131 AGENCOURT
5	1340	92.4	1004	14	CD515356	CD515356 AGENCOURT
6	1292	89.1	1129	12	BM802792	BM802792 AGENCOURT
7	1279	88.2	871	14	CA488274	CA488274 AGENCOURT
8	1205	83.1	1988	11	BC023725	BC023725 Mus muscu
9	1205	83.1	3110	11	AK031542	AK031542 Mus muscu
10	1205	83.1	3372	11	BC046553	BC046553 Mus muscu
11	1205	83.1	4323	11	AK031099	AK031099 Mus muscu
12	1205	83.1	4463	11	AK083669	AK083669 Mus muscu
13	1165.5	80.4	852	14	CD516283	CD516283 AGENCOURT
14	1136.5	78.4	932	14	CA487981	CA487981 AGENCOURT
15	1127	77.7	685	9	AU123871	AU123871 AU123871
16	1092	75.3	653	12	BM721340	BM721340 UI-E-E01-
17	1084	74.8	902	14	CA48868	CA48868 AGENCOURT
18	1070	73.8	791	12	BI154872	BI154872 602902857
19	1035	71.4	887	14	CA980253	CA980253 AGENCOURT
20	1032	71.2	943	13	BQ958632	BQ958632 AGENCOURT
21	966	66.6	919	13	B0845391	B0845391 AGENCOURT
22	954	65.8	537	12	BM793744	BM793744 K-EST0074
23	936	64.6	813	12	BI557797	BI557797 603236977
24	933	64.3	899	14	CD558386	CD558386 AGENCOURT
25	925	63.8	855	13	BU594980	BU594980 AGENCOURT
26	915	63.1	717	14	CB598701	CB598701 AGENCOURT
27	899.5	62.0	916	13	BU845369	BU845369 AGENCOURT
28	898	61.9	906	14	CA454570	CA454570 AGENCOURT
29	897	61.9	547	12	BM787824	BM787824 K-EST0066
30	896	61.8	491	12	BM790293	BM790293 K-EST0070
31	889	61.3	998	12	BI649877	BI649877 603296516
32	880.5	60.7	787	12	BI156730	BI156730 602922150
33	880.5	60.7	844	12	BI158626	BI158626 602922395
34	880	60.7	609	9	AW701942	AW701942 ug93b02.y
35	874	60.3	621	10	BG283493	BG283493 602407782
36	861.5	59.4	849	14	CA489799	CA489799 AGENCOURT
37	860	59.3	786	12	BI155788	BI155788 602904360
38	858	59.2	887	13	BQ717097	BQ717097 AGENCOURT
39	843.5	58.2	649	9	AA496412	AA496412 zv37a02.r
40	829.5	57.2	629	10	BE374596	BE374596 601225384
41	829	57.2	533	14	CB437620	CB437620 685491.MA
42	829	57.2	552	10	BG277542	BG277542 ux45d07.y
43	817	56.3	518	10	BF039952	BF039952 BP250015A
44	809	55.8	575	12	BG70469	BG70469 602838347
45	804.5	55.5	678	12	BI555157	BI555157 603236396

ALIGNMENTS

RESULT 1
CA455074
LOCUS CA455074 885 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT_10735919 MAPCL Homo sapiens cdna clone IMAGE:6722585 5',
mRNA sequence.
ACCESSION CA455074
VERSION CA455074.1 GI:24905427
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 885)

AUTHORS NTH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@email.nih.gov
 Tissue Procurement: Kristi A. Eglund, Ira Pastan
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
 Plate: L1AMI4285 row: 1 column: 17
 High quality sequence stop: 717.

FEATURES
 source
 1..885
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6722585"
 /cell_line="2R-75-1, MCF7, SK-BR-3, MDA-MB-231, HERT-HME1
 , LNCap"
 /lab_host="EMDH10B"
 /clone_lib="MAPCL"
 /note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
 Subtracted with brain, liver, lung, kidney and muscle.
 Directionally cloned. Priming method: oligo-dr. Average
 insert size: 1800 bp. Library amplification: 26,000 fold.
 Kristi A. Eglund, James J. Vincent, Robert Strausberg,
 Bungkok Lee & Ira Pastan: Discovery of new breast
 cancer genes encoding membrane and secreted proteins.
 Manuscript submitted."

BASE COUNT 178 a 297 c 261 g 149 t

ORIGIN

Alignment Scores:

Pred. No.: 1.61e-79 Length: 885
 Score: 1426.00 Matches: 264
 Percent Similarity: 99.62% Conservative: 0
 Best Local Similarity: 99.62% Mismatches: 0
 Query Match: 98.34% Indels: 1
 DB: 14 Gaps: 0

SEQ4 (1-266) x CA455074 (1-885)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 Db 1 CAGAAATGAGGACTTGGCCCGACGAGTCCCTTGGACAGCACCTTCTACCGCTCAGTCTG 60
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
 Db 61 GAGGACGATCACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 120
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
 Db 121 TTCTTCTGTCAGACCTCGCCCGCGCGCTGGGGGATGGTCCACACAGCAGCCGACG 180
 Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
 Db 181 TCATCTACAGAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 240
 Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
 Db 241 GCCCCAGGCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGATATTGTGTTGAC 300
 Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisaspProSerProLeu 120
 Db 301 CTGGGAATGGGGCAGCAGCAGGGGCTGCAAGGCTCCCCACATGACCCACCCCTCTA 360
 Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
 Db 361 CAGCGGTACAGTACGAGNCCCAAGTACCCCTGCCCTCTGAGACTGATGGCTACGTTGCC 420

Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
 Db 421 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACAGCCAGATGTTCGCCCCACGCC 480
 Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
 Db 481 CCTTCGCCCGAGAGGGCCCTCTGCCTGTGCGCCGACCTGCTGCTGCTGCTGCTGCTG 540
 Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
 Db 541 CCCAAGACTCTCTCCCGAGGAGAAATGGGGTGTCAAGACGCTTTTGTCTTTGGGGGT 600
 Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
 Db 601 GCGGTGGAGAACCCCGAGTACTTGCACACCCAGGAGGAGCTGCCCTCAGCCGCCACCT 660
 Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
 Db 661 CCTCCTGCCTTCAGCCCGACCTTCGACAACTCTATTACTTGGACGAGCAGCCACCCAG 720
 Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
 Db 721 CGGGGGGCTCCACCCAGCACCTTCAAGGGGACACTACGCGAGAACCCAGAGTACCTG 780
 Qy 261 -GlyLeuAspVal 264
 Db 781 GGCTCTGGAGGTG 793

RESULT 2
 AF318349
 LOCUS Homo sapiens pp3659 mRNA, complete cds.
 DEFINITION AF318349
 ACCESSION AF318349.1 GI:18027789
 VERSION
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4715)
 AUTHORS Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T.,
 Wan, D.F. and Gu, J.R.
 TITLE Novel human cDNA clones with function of inhibiting cancer cell
 growth
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 4715)
 AUTHORS Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T.,
 Wan, D.F. and Gu, J.R.
 TITLE Direct Submission
 JOURNAL Submitted (02-NOV-2000) National Laboratory For Oncogenes & Related
 Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai
 200032, P. R. China

FEATURES
 source
 1..4715
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..4715
 /gene="pp3659"
 2950..4188
 /gene="pp3659"
 /codon_start=1
 /product="unknown"
 /protein_id="AAL55856.1"
 /db_xref="GI:18027790"
 /translation="MALESILRRFRTHQSDVWSYGVTVWELMTFGAKPYDGIPIAREIP
 DLLEKERLPQPPDICTDVMIMVKMIDSECRPRELVSSEFSRNARDPQRFVVIQ
 NEDLGASPLDSTFYRSLLEDMDGLVDAAEYLVVQQGFCDPAPAGGMVHRRH
 SSSTRSGGDLTLGLPSEEEAPSLAPSEGAGSDVFDGLDGMGAAGLQSLPTDP
 SPQRYSEDPVPLPSETDGYVAPLTCSPQEVYNQPDVRFQPPREGPLPAARAP
 ATLERPKTILSPCKNGVVKDVFAGGAVENPEYLTPOGGALSPDLLPSAPSTST
 GTRTHSGGLHPAPSKGLHLKQRTQSTWVWTCQCEPQGVRRSPDVSSSGSREGLTSGI
 KRWEGPPTTSRGTCHARNLS"


```

BASE COUNT      969 a   1446 c   1347 g   953 t
ORIGIN
Alignment Scores:
Pred. No.:      1,21e-78      Length:      4715
Score:          1423.00      Matches:      265
Percent Similarity: 99.62%      Conservative: 0
Best Local Similarity: 99.62%      Mismatches: 1
Query Match:    98.14%      Indels:      1
DB:             11      Gaps:      0

SEQ4 (1-266) x AF318349 (1-4715)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3253 CAGATGAGGACTGGGGCCAGCCAGTCCCTTGACAGCAGCCTCTACCGCTCACTGCTG 3312
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3313 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3372
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60
Db 3373 TTCCTCTGTCAGACCTGCGCCCGGGCGCTGGGGCGATGGTCCACAGCAGCCGCGAGC 3432
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3433 TCATCTACGAGAGTGGGGTGGGACCTGACATGAGGCTGGAGCCCTCTGAAGAGGAG 3492
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3493 GCCCCAGGCTCCACATGGCACCTCCGAAGGGGCTGGCTCCGATGTATTGTGTTGTCAC 3552
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3553 CTGGGAATGGGGGAGCCAGGCGGCTGCAAAAGCCCTCCACACATGACCCCGAGCCCTCTA 3612
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3613 CAGGGTACATGAGGACCCACAGTACCTCGCCCTGAGACTGATGGCTAGCTTGC 3672
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3673 CCCCTGACCTGCAGCCCCAGCCCTGAATATGTGAACACGACGAGTGTTCGCCCCAGGCC 3732
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3733 CCTTCGCCCGGAGAGGGCCCTCTGCTGCTGCCCGACCTGTGTGTCACACTCTGGAAGG 3792
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValLysAspValPheAlaPheGlyGly 200
Db 3793 CCCAAGACTCTCTCCCGAGGAGAGATGGGGTGTGCAAAAGAGCTTTTTCCTCTGGGGGT 3852
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3853 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGTGC-CCTCAGCCCCACCCCT 3911
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3912 CCTCCTGCTTTCAGCCCGACCTTGCACAACCTCTATTACTGGGACGAGGCCACCCAGAG 3971
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3972 CGGGGGGCTCCACCCAGACCCCTCAAGGGGACACCTACGGGAGAGAACCCAGAGTACCTG 4031
Qy 261 GlyLeuAspValProVal 266
Db 4032 GGTCTGGACGTGCCAGTG 4049

RESULT 3
CA455141      893 bp      mRNA      linear      EST 12-NOV-2002
LOCUS
DEFINITION  AGENCOUNT_10735980 MAPcL Homo sapiens cDNA clone IMAGE:6722663 5',

```

```

mRNA sequence.
ACCESSION      CA455141
VERSION        CA455141.1
KEYWORDS       GI:24905561
SOURCE         EST.
ORGANISM       Homo sapiens (human)
REFERENCE      1 (bases 1 to 893)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished
COMMENT        Contact: Robert Strausberg, Ph.D.
                Email: cgapbs@mail.nih.gov
                Tissue Procurement: Kristi A. Eglund, Ira Pastan
                cDNA Library Preparation: Invitrogen Corp
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Agencourt Bioscience Corporation
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLAM14285 row: 1 column: 23
                High quality sequence stop: 682.
                Location/Qualifiers
                    1..893
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:6722663"
                        /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1
                        , LNCap"
                        /lab_host="EMDH108"
                        /clone_lib="MAPcL"
                        /note="Vector: PCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
                        Subtracted with brain, liver, lung, kidney and muscle.
                        Directionally cloned. Priming method: oligo-dT. Average
                        insert size: 1800 bp. Library amplification: 26,000 fold.
                        Kristi A. Eglund, James J. Vincent, Robert Strausberg,
                        Bungkok Lee & Ira Pastan: Discovery of new breast
                        cancer genes encoding membrane and secreted proteins.
                        Manuscript submitted."
BASE COUNT      175 a   302 c   261 g   155 t
ORIGIN
Alignment Scores:
Pred. No.:      7,21e-78      Length:      893
Score:          1399.50      Matches:      261
Percent Similarity: 98.13%      Conservative: 1
Best Local Similarity: 97.75%      Mismatches: 3
Query Match:    96.52%      Indels:      2
DB:             14      Gaps:      1

SEQ4 (1-266) x CA455141 (1-893)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 67 CAGATGAGGACTGGGGCCAGCCAGTCCCTTGAGCAGCAGCCTTACCGCTCACTGCTG 126
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 127 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGTTACCCAGCAGGCG 186
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60
Db 187 TTCCTCTGTCAGACCTGCGCCCGGGCGCTGGGGCGATGGTCCACAGCAGCCGCGAGC 246
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 247 TCATCTACGAGAGTGGGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 306
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 307 GCCCCAGGCTCTCCACTGGCACCCCTCCGAAGGGGCTGCTCCGATGTATTGTGTTGTCAC 366

```

```

101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
    |||||
367 CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCACACATGACCCAGCCCTCTA 426
    |||||
121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
    |||||
427 CACGGGTACAGTGGAGGACCCAGACATACCCCTGCTGAGACTGATGGCTACGTGGC 486
    |||||
141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
    |||||
487 CCCTTGACCTGCAGCCGCCAGCTGATATGTGAACAGCCAGATGTCGGCCCGAGCC 546
    |||||
161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
    |||||
547 CCTTCGCCCGCAGAGGGCCCTGCTGCTGCTGCCGACCTGCTGGTGCACCTCTGGAAGG 606
    |||||
181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
    |||||
607 CCCAAGACTCTCTCCCGAGGAAGATGGGGTCTGCAAAAGACGTTTTTGGCTTTGGGGT 666
    |||||
201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
    |||||
667 GCGGTGGAGAACCCCGAGTACTTGACACCCAGGAGAGCTGCCCTCAGCCCCACCT 726
    |||||
221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
    |||||
727 CCTCCTGCCCTTACGCCAGCCTTCGACACCTCTATTACTGGACAGGACCCACCCAGAA 786
    |||||
241 -ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLe 260
    |||||
787 GCGGGGGGCTCCACCCAGCACCTTCAAAAGGGACACCTACGGCAGAGAACCACCCAGATACC 846
    |||||
260 uGlyLeu---AspValPro 265
    |||||
847 TGGGGTCTGGGACGTGCC 865
    |||||

RESULT 4
CA454131
LOCUS
DEFINITION
AGENCOURT_10738550 MAPcL Homo sapiens cDNA clone IMAGE:6718792 5',
mRNA sequence.
CA454131
VERSION
KEYWORDS
SOURCE
EST.
GI:24903569
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 894)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M14275 row: k column: 16
High quality sequence stop: 636.
Location/Qualifiers
1..894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6718792"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HMEL
, LNCaP"
/lab_host="EMDH10B"
FEATURES
source

```

```

/clone_lib="MAPcL"
/notes="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."
BASE COUNT 180 a 303 c 260 g 151 t
ORIGIN

Alignment Scores:
Pred. No.: 4.86e-75 Length: 894
Score: 1354.00 Matches: 262
Percent Similarity: 96.69% Conservatives: 1
Best Local Similarity: 96.32% Mismatches: 3
Query Match: 93.38% Indels: 6
DB: 14 Gaps: 0

SEQ4 (1-266) x CA454131 (1-894)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 24 CAGAATGAGGACTTGGGCCCGCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 83
Qy 21 GluAspAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGly 40
Db 84 GAGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 143
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHISArgHisArgSer 60
Db 144 TTCTTCTGCACACCTGCCCCGGCGCTGGGGCATGCTGCACACAGCAGCCGCGAGC 203
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 204 TCATCTACCGAGGAGTGGCGTGGGACCTAGGGCTGGAGCCCTCTGAAGAGGAG 263
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 264 GCCCCCAGGTCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGATTTGATGTGAC 323
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 324 CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCCCACACATGACCCAGCCCTCTA 383
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 384 CACGGGTACAGTGGAGGACCCAGTACCCCTGCTGAGACTGATGGCTACGTGGCC 443
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 444 CCCTTGACCTGCAGCCCGCCAGCTGATATGTGAACAGCCAGATGTCGGCCCGCAGCC 503
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 504 CCTTCGCCCGCAGAGGGCCCTGCTGCTGCTGCCGACCTGCTGGTGCACCTCTGGAAGG 563
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 564 CCCAAGACTCTCTCCCGAGGAAGATGGGGTCTGCAAAAGACGTTTTTGGCTTTGGGGT 623
Qy 201 AlaValGluAsnProGluTyrLeuThrProGln-GlyGlyAlaAlaProGlnProHisPr 220
Db 624 GCGGTGGAGAACCCCGAGTACTTGACACCCCGGGGAGGAGCTGCCCTCAGCCCCACCC 683
Qy 220 OProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrp-AspGlnAspProPro 240
Db 684 TCCTCTGCTCCTCAGCCCGACCTTCGACAACTCTATTACTGGGACCCAGGACCCCGAGG 743
Qy 240 LuArgGlyAla--ProProSerThrPheLysGlyThrProThrAlaGluAsnProGlu-T 259
Db 744 AGCGGGGGGGCTCCCGACCCAGCACCTTTCAAGAGGACACCTACGCGCAGAGAACCAGAGT 803

```

[illegible][illegible]

Db	663	ACCTTCAAAGGACACCTACGGCAGAGAACCCAGAGTACCTACCTGGGGTCTGGAGCTGCCAGT	722
QY	266	1	266
Db	723	G	723
RESULT 7	CA488274	871 bp	linear
LOCUS	CA488274	MAPCL Homo sapiens	EST 14-NOV-2002
DEFINITION	AGENCOURT_10808010	MAPCL Homo sapiens	cdna clone IMAGE:6720019 5',
ACCESSION	CA488274	1	GI:24949701
VERSION	CA488274.1		
KEYWORDS	EST		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 871)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Kristi A. Eglund, Ira Pastan CDNA Library Preparation: Invitrogen Corp CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM14278 row: n column: 19 High quality sequence stop: 693.		
FEATURES	location/Qualifiers		
	1..871		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6720019"		
	/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HTERT-HME1", LNCap"		
	/lab_host="EMDH10B"		
	/clone_lib="MAPCL"		
	/note="Vector: pCMV-SPORT6; Site_1: EcorV; Site_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dr. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan. Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."		
BASE COUNT	177 a	296 c	144 t
ORIGIN			2 others
Alignment Scores:			
Pred. No.:	2,18e-70	Length:	871
Score:	1279.00	Matches:	236
Percent Similarity:	99.16%	Conservative:	0
Best Local Similarity:	99.16%	Mismatches:	1
Query Match:	88.21%	Indels:	1
DB:	14	Gaps:	0
SEQ4 (1-266) x CA488274 (1-871)			
QY	30	AspAlaGluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGly	49
Db	3	GATGTGAGGAGTATCGTACCCAGCAGGGCTCTCTCTGTCCAGACCCTGCCCGGGC	62
QY	50	AlaGlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAsp	69
Db	63	GCTGGGGGCATGGTCCACACAGCAGCCGAGCTCATCTACAGGAGTGGCGTGGGAC	122
QY	70	LeuThrLeuGlyLeuGluProSerGluGluAlaProArqSerProLeuAlaProSer	89

```

|||||
123 CTGACACTAGGCGTGGAGCCCTCTCAAGAGAGGCCCCCAGGCTCCACCTGCACCCCTCC 182
QY 90 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeu 109
|||||
183 GAAGGGGTGGCTCCGATGATTTGATGCTGACCTGGGAATGGGGCCAGCCAAAGGGGCTG 242
QY 110 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 129
|||||
243 CAAGCCCTCCACACATGACCCAGCCCTCTACAGCGGTACAGGTGAGGACCCACAGTA 302
QY 130 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlu 149
|||||
303 CCCCTGCCCTCTGAGACTGATGGCTACGTTGCCCTGACCTGACCTGCAGCCGCCAGCTGAA 362
QY 150 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 169
|||||
363 TATGTGAACACAGCCAGATGTTGGGCCCCAGCCCTGCGCCGAGAGGGCCCTCGCT 422
QY 170 AlaAlaArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsn 189
|||||
423 GCTGCCACACCTGCTGGTCCACTCTGGAAGGCCACAGACTCTCTCCCGAGGGAAGAT 482
QY 190 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 209
|||||
483 GGGGTGCTCAAGACGCTTTTGCCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACA 542
QY 210 ProGlnGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 229
|||||
543 CCCAGGGAGGAGCTGCCCTCAGCCACCCCTCTCTGCTTCCAGCCAGCCCTTCGAC 602
QY 230 AsnLeuTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyr 249
|||||
603 AACCTCTATTACTGGACACAGGACCCACAGAGCGGGGGCTCCACCCAGCACCTTCAA 662
QY 250 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 266
|||||
663 GGGACACCTACGGCAGAGAACCCAGACTACCTGGNGTCTGGAGCTGCCAGTG 714

RESULT 8
BC023725 LOCUS
DEFINITION
  Mus musculus, 1988 bp mRNA linear HTC 04-MAR-2003
  oncogene homologue 2 (neuro/glioblastoma derived oncogene homolog),
  clone IMAGE:5347334, mRNA.
ACCESSION BC023725
VERSION BC023725.1 GI:23959125.
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1988)
Strausberg,R.
Direct Submission
Submitted (05-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-re@mail.nih.gov
Tissue procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,

```

```

Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAP Plate: 55 Row: 1 Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: retained intron.

FEATURES
    source
        1..1988
            Location/Qualifiers
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:5347334"
                /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
                ductal carcinoma. 5 month old virgin mouse."
                /clone_lib="NCI_CGAP_Mam6"
                /lab_host="DH10B"
                /note="Vector: pCMV-SPORT6"

BASE COUNT 444 a 580 c 532 g 432 t
ORIGIN

Alignment Scores:
Pred. No.: 1..89e-65 Length: 1988
Score: 1205.00 Matches: 223
Percent Similarity: 87.97% Conservative: 11
Best Local Similarity: 83.83% Mismatches: 32
Query Match: 83.10% Indels: 0
DB: 11 Gaps: 0

SEQ4 (1-266) x BC023725 (1-1988)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
123 CAGAACGAGGACTTAGGCCCTCCAGCCCCATGGAGCAGCACCTTCTACCGTTCTACCTG 182
QY 21 GluAspAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
|||||
183 GAGGATGATGACATGGGGAGCTGCTGATGCTGAAGACTGCTGTTACCCAGCAGGGA 242
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60
|||||
243 TTCTTCTCCAGACCTGCCCTAGGTACTGGGAGCAGACAGCCAGCAGACACCGCAGC 302
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGluProSerGluGluGlu 80
|||||
303 TCGTCGGCCAGAGTGGCGGTGGTGAAGTGAAGTGGGCTGGAGCCCTCGGAGGAAGAG 362
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
|||||
363 CCCCCACATCTCCACTGGCTCCCTCCGAGGGGCTGGCTCCGATGCTGTTGATGGTAC 422
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
|||||
423 CTGGCAGTGGGGGTAAACCAAGGACTGACAGAGCCTCTCTCCACATGACCTACGCCCTCA 482
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
|||||
483 CAGCGGTACAGTGAAGATCCCATATTACTCTGCCCCCGAGACTGATGGCTTACGTTGCT 542
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
|||||
543 CCCCTGGGCTGCAGGCCGCCAGCCAGATGTAACACAGAGAGGTTCCGGCCTCAGTCT 602
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
|||||
603 CCCTTGACCCAGAGGGTCTCTCCGCTCCCATCCGACCTGCTGGTGTCTACTCTAGAAAGA 662

```



```

SSSVARCPSPGVKPDLSYMPIWKYPDEEGICQPCPINCHSCVLDLDERGCPAEORASP
VTFIATVGVLLFLIIIVVIGILIKRRRQIRKYITMRRLQETELVPLTPSPGAPVN
QAOMILAKETELRLKLVKGAGFTVYKGIWIPDGENVKIIPALIKVLRETPSPKANKE
ILDEAYVWAGVSPVSRLLGICLTSTVQLTQIMPYGCCLDHVREHGRGRGSLQDLN
WCQVIAKMSILEEVRVLVRLDAAKNVLPKPNHVKIINDFGLARLLDIDETEYHADGG
KVPITKMALESILRRRFTHQSDVMSYGVTVMLMTFGAKPYDGIPIPARIPDLLEKGER
LPQPDICTIDVYIMVKWMDISECRPRFELVSEFRMARDPQRFVYQVIONEDLGPS
PMDSTFYRSLLEDODDDELVAEYLVPOQGFSPDPALGTGTSTAHRRHRSRSARSGG
GELTGLSPSEEEPRPLAPSEAGSDVFDGLAVGTYKGLQSLSPHDLSPLORYSE
DPTLPLPETGCVYAPLACSPQPEYVNOPEVRPQSLPLPEGPPPIRIRPAGATLERPKT
LSPGKNGVKDVFAFGAVENPEYLAPRAGTASQPHSPAFSPAFDNLIIYWDNSSEQ
GPPSTFGTGPTAENPEYLGIDVPV*
BASE COUNT      677 a      886 c      891 g      656 t
ORIGIN

```

```

Alignment Scores:
Pred. No.:      2,89e-65      Length:      3110
Score:          1205.00      Matches:      223
Percent Similarity: 87.97%      Conservative: 11
Best Local Similarity: 83.83%      Mismatches: 32
Query Match:      83.10%      Indels:      0
DB:              11          Gaps:      0

```

SEQ4 (1-266) x AK031542 (1-3110)

```

QY      1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB      1848 CAGACGAGGACTTAGGCCCTCCAGCCCCATGGACAGACCTTCTACGTTCACTGCTG 1907

QY      21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGly 40
DB      1908 GAGGATGATGACATGGGGAGCTGCTGATGCTGAAGAGTACCTGGTACCCAGCAGGGA 1967

QY      41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60
DB      1968 TTCTTCTCCACAGACCTGCGCCTAGGTACTGGAGCACAGCCACCGCAGACACCGCAGC 2027

QY      61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB      2028 TCCTCGCCGAGGAGTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2087

QY      81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB      2088 CCCCCAGATCTCCACTGGCTCCCTCCGAAAGGGGCTGCTCGATGTGTTGATGCTGATGCTG 2147

QY      101 LeuGlyMetGlyAlaAlaGlySerLeuGlnSerLeuProThrHisAspProSerProLeu 120
DB      2148 CTGGCAGTGGGGGGTAAACCAAGGACTGCAGAGCCCTCTCTCCACATGACCTCAGCCCTCTA 2207

QY      121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
DB      2208 CAGCGGTACAGTAGAGATCCACATTAATCTCTGCCCCCGAGAGCTGATGGCTACGTTGCT 2267

QY      141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
DB      2268 CCCCCTGGCTGCAGCCCCCAGCCGAGTATGTGAACACGACGAGAGGTTCCGGCTCAGTCT 2327

QY      161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
DB      2328 CCCCCTGACCCAGAGGGGCTCCCGCTCCCATCCGACCTGCTGGTGTCTACTCTAGAAAGA 2387

QY      181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
DB      2388 CCCAAGACTCTCTCTCCGGGAAATGGGGTGTGAAGAGCTTTTTCCTTTGGGGGT 2447

QY      201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
DB      2448 GCTGTGGAGAACCTGAATACTTAGCACCCAGCAGGAGGACTGCTCTCAGCCCCACCT 2507

QY      221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTyrAspGlnAspProGlu 240
DB      2508 TCTCTGCTTCCAGGCCCGCTTTTGACAACTCTATTACTGGGACCAAGAACTCATCGGAG 2567

```

```

QY      241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
DB      2568 CAGGTCCTCCACCAAGTACCTTTGAAGGGACCCCTGCAGAGACCTCGAGTACCTA 2627

QY      261 GlyLeuAspValProVal 266
DB      2628 GGCGTGGATGTGCCAGTA 2645

RESULT 10
BC046553      3372 bp      mRNA      linear      HTC 03-FEB-2003
LOCUS      Mus musculus, clone IMAGE:5340777, mRNA.
DEFINITION      BC046553
ACCESSION      BC046553.1 GI:28196923
VERSION      HTC.
KEYWORDS      Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM      Mus musculus
REFERENCE      1 (bases 1 to 3372)
AUTHORS      Strausberg, R.
TITLES      Direct Submission
JOURNAL      Submitted (03-FEB-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 92 Row: n Column: 11
This clone has the following problem: no 5' EST match.

FEATURES
source
location/Qualifiers
1..3372
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5340777"
/tissue_type="Mammary tumor. Brcal-/fl; MMTV-Cre model. 10
months old, gross tissue."
/clone_lib="NCI CGAP_Mam3"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

BASE COUNT      754 a      944 c      926 g      746 t
ORIGIN

Alignment Scores:
Pred. No.:      3.12e-65      Length:      3372
Score:          1205.00      Matches:      223
Percent Similarity: 87.97%      Conservative: 11
Best Local Similarity: 83.83%      Mismatches: 32
Query Match:      83.10%      Indels:      0
DB:              11          Gaps:      0

SEQ4 (1-266) x BC046553 (1-3372)
QY      1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB      1505 CAGACGAGGACTTAGGCCCTCCAGCCCCATGGACAGACCTTCTACGTTCACTGCTG 1564

```


Qy	21	GlulaspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly	40
Ds	1565	GAGGATGATGACATGGGGAGTGGTGTGATGCTGAAGAGTACCTGGTACCCAGCAGGA	1624
Qy	41	PhePheCysProaspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer	60
Ds	1625	TTCTTCTCCCGACACCTGGCTAGGTACTGGAGCAGACGCCACCCGACACCGCAGC	1684
Qy	61	SerSerThrArgSerGlyGlyAspLeuThrLeuGluLeuGluProSerGluGluGlu	80
Ds	1685	TCGTGGCCAGGAGTGGCGTGGAGCTGACATGGGCTGGAGCCCTCGGAAGAAGAG	1744
Qy	81	AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp	100
Ds	1745	CCCCCAGATCTCCAGTGGCTCCCTCCGAAGGGCTGGCTCCGATGTGTGTGATGGTAC	1804
Qy	101	LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu	120
Ds	1805	CTGGCAGTGGGGTAAACCAAGGACTGCAGAGCTCTCTCCACATGACCTCAGCCCTCTA	1864
Qy	121	GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla	140
Ds	1865	CAGCGTACAGTGGAGTCCCATTTACCTCTGCCCCCGAGACTGATGGCTACGTTGCT	1924
Qy	141	ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro	160
Ds	1925	CCCTGGCTGCAGCCCCAGCCGAGTATGTGAACCCAGCCAGAGGTTCGCGCTCAGTCT	1984
Qy	161	ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg	180
Ds	1985	CCCTTGACCCAGAGGCTCTCCGCTCCCATCCGACTCGCTGCTACTCTAGAAAGA	2044
Qy	181	ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly	200
Ds	2045	CCCAAGACTCTCTCTGGGAAATGGGGTGTCAAAGACGTTTGTGCTTGGGGT	2104
Qy	201	AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro	220
Ds	2105	GCTGTGAGAACCTGTAATCTTAGCACCCAGAGCAGGACCTGCTCTCAGCCGCCACCT	2164
Qy	221	ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu	240
Ds	2165	TCCTCTCCCTCAGCCCGCTTGAACCTCTATTACTGGGACCAAGAACTCATCGGAG	2224
Qy	241	ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu	260
Ds	2225	CAGGCTCTCCACCAAGTACCTTTGAAGGGACCCCTCCTGCAGAGAACCTTGAGTACCTA	2284
Qy	261	GlyLeuAspValProVal	266
Ds	2285	GGCCTGGATGTCAGTA	2302
RESULT	11		
AK031099			
LOCUS			
DEFINITION		4323 bp mRNA linear HTC 05-DEC-2002	
		Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length	
		enriched library, clone:5930404M10 product:v-erb-b2 erythroblastic	
		leukemia viral oncogene homolog 2, neuro/glioblastoma derived	
		oncogene homolog (avian), full insert sequence.	
ACCESSION	AK031099		
VERSION	AK031099.1	GI:26082143	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
JOURNAL	1		
MEDLINE	Carninci,P. and Hayashizaki,Y.		
PUBMED	High-efficiency full-length cDNA cloning		
	Meth. Enzymol. 303, 19-44 (1999)		
	99279253		
	10349636		
REFERENCE			
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,		
TITLE	Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to		
MEDLINE	Prepare full-length cDNA libraries for rapid discovery of new genes		
PUBMED	Genome Res. 10 (10), 1617-1630 (2000)		
	20499374		
	11042159		
REFERENCE			
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,		
	Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,		
	Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,		
	Yamanoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,		
	Fujiwaka,S., Inoue,K., Toqawa,Y., Izawa,M., Ohara,E., Watahiki,M.,		
	Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,		
	Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.		
	RIKEN integrated sequence analysis (RISA) system--384-format		
	sequencing pipeline with 384 multicapillary sequencer		
	Genome Res. 10 (11), 1757-1771 (2000)		
	20530913		
	11076861		
REFERENCE			
AUTHORS	Kawai,J., Shingawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,		
	Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,		
	Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,		
	Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,		
	Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,		
	Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,		
	Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,S.,		
	Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,		
	Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,		
	Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,		
	Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,		
	Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,		
	Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,		
	Marchionni,L., Mashima,J., Mazzarelli,J., Mommaerts,P., Nordone,P.,		
	Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,		
	Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,		
	Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,		
	Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawajiri,H., Kohtsuki,S.		
	and Hayashizaki,Y.		
	Functional annotation of a full-length mouse cDNA collection		
	Nature 409 (6821), 685-690 (2001)		
	21085660		
	11217851		
REFERENCE			
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research		
	Group Phase I & II Team.		
	Analysis of the mouse transcriptome based on functional annotation		
	of 60,770 full-length cDNAs		
	Nature 420, 563-573 (2002)		
	6 (bases 1 to 4323)		
	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,		
	Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,		
	Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,		
	Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,		
	Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,		
	Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,		
	Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,		
	Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,		
	Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T.,		
	Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,		
	Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,		
	Muramatsu,M. and Hayashizaki,Y.		
	Direct Submission		
	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of		
	Physical and Chemical Research (RIKEN), Laboratory for Genome		
	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),		
	RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,		
	kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp,		
	URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,		
	Fax:81-45-503-9216)		
	cDNA library was prepared and sequenced in Mouse Genome		

Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES

source
1. .4323
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:5930404N10"
/db_xref="taxon:10090"
/clone="5930404N10"
/tissue_type="forelimb"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
48..3818
/note="putative
v-erb-b2 erythroblastic leukemia viral oncogene homolog 2,
neuro/glioblastoma derived oncogene homolog (avian)
(MGD|MG1:95410, GBU71126, evidence: BLASTN, 99%,
match=449)"
misc_feature 933 a 1247 c 1226 g 917 t
BASE COUNT 933 a 1247 c 1226 g 917 t
ORIGIN

Alignment Scores:

Score: 3.95e-65 Length: 4323
Pred. No.: 1205.00 Matches: 223
Percent Similarity: 87.97% Conservative: 11
Best Local Similarity: 83.83% Mismatches: 32
Query Match: 83.10% Indels: 0
DB: 11 Gaps: 0

SEQ4 (1-266) x AK031099 (1-4323)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3018 CAGACGAGGACTTAGGCCCCCTCCAGCCCCATGGAGCAGCCTTCTACGTTCTGCTG 3077
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGly 40
Db 3078 GAGGATGATGACATGGGGGAGCTGCTGATGCTGAAGAGTACCTGTTACCCAGCAGGGA 3137
QY 41 PhePheCysProAspProAlaGlyGlyMetValHisArgHisArgSer 60
Db 3138 TTCTTCTCCAGAGCCCTGCTAGGTACTGGGAGCAGACCCAGCAGACCGCAGC 3197
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 80
Db 3198 TCGTCGGCCAGAGTGCGGTGTGCTGACACTGGGCTGGAGCCCTCGGAAGAAGAG 3257
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3258 CCCCCAGACTCTCCACTGGCTCCCTCCGAAAGGGGCTGCTCGATGTGTTGATGGTGAC 3317
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3318 CTGGCAGTGGGGGTAAACAAAGAGCTGCAGAGCCCTCTTCCACATGACCTCAGCCCTCTA 3377
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3378 CAGCGGTACAGTAGAGATCCACATTAACCTCTGCCCCCGAGACTGATGGCTACGTTGCT 3437
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3438 CCCCCTGGCTGTCAGCCCCAGCCCGAGTATGTGAACACAGCCAGAGAGTTCCGGCTCAGTCT 3497
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3498 CCCCCTGACCCAGAGGGTCCCTCCGCTCCCATCCGACCTGCTGGTGGTACTCTAGAAAGA 3557

QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3558 CCCAAGACTCTCTCTCCGGGAAATGGGTGTCAAGACGTTTTTCCCTTGGGGCT 3617
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 3618 GCTGGGAGAACCCCTGAATACCTAGCCAGCAGCAGCAGCTCCCTCTCAGCCCCACCT 3677
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTyrAspGlnAspProProGlu 240
Db 3678 TCTCTCTGCTTCAGCCCGCCCTTTCACCAACCTCTATTACTGGGACGACGACATCATCGGAG 3737
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3738 CAGGGTCTCTCCACCAAGTACCTTTGAAGGACCCCTCAGAGAACCTGAGTACCTA 3797
QY 261 GlyLeuAspValProVal 266
Db 3798 GGCCTGGATGTGCCAGTA 3815
RESULT 12
AK083669
LOCUS
DEFINITION
Mus musculus 9 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:D030063B12 product:v-erb-b2 erythroblastic
leukemia viral oncogene homolog 2, neuro/glioblastoma derived
oncogene homolog (avian), full insert sequence.
ACCESSION AK083669
VERSION AK083669.1 GI:26101404
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, K.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Stauble, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,

Fletcher,C., Fujita,M., Gariboldi,M., Gustinich,S., Hill,D., Hofmann,M., Hume,D.A., Kaniya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S., and Hayashizaki,Y.

Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851

5
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 4463)

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.

FEATURES
 source
 Location/Qualifiers
 1..4463
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM,DB:D030063B12"
 /db_xref="taxon:10090"
 /clone="D030063B12"
 /tissue_type="whole body"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="9 days embryo"

misc_feature
 1..4463
 /note="v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian) (MGDI|MG1:95410, GB|U71126, evidence: BLASTN, 99%, match=449)";

BASE COUNT
 950 a 1290 c 1270 g 953 t

ORIGIN
 Alignment Scores:
 Pred. No.: 4.07e-65 Length: 4463
 Score: 1205.00 Matches: 223
 Percent Similarity: 87.97% Conservative: 11
 Best Local Similarity: 83.83% Mismatches: 32
 Query Match: 83.10% Indels: 0
 DB: 11 Gaps: 0

SEQ4 (1-266) x AK083669 (1-4463)

QY 1 GLAAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 |||||
 Db 3100 CAGAACGAGGACCTAGGCGCCCTCCAGCCCATGAGACACCTTCTACCGTTCACTGCTG 3159
 |||||

QY 21 GLAAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
 |||||
 Db 3160 GAGGATGATCATGGGGAGCTGGTGTGCTGAACAGTACCTGGTACCCAGCAGGGA 3219
 |||||

QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
 |||||
 Db 3220 TTCTTCTCCAGACCTGCTGCTAGTACTGGGAGCAGCCACCCAGCAGACCCGAGC 3279
 |||||

QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
 |||||
 Db 3280 TCGTGGCCAGGAGTGGCGGTGGAGCTGACACTGGGCTGGAGCCCTCGGAAGAAGAG 3339
 |||||

QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
 |||||
 Db 3340 CCCCCAGATCTCCACTGGCTCCCTCCGAAAGGGCTGGCTCCGATGTTGATGGTAC 3399
 |||||

QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
 |||||
 Db 3400 CTGCGAGTGGGGGTAAACAAAGGAGCTGCAGAGCCTCTCCACATGACCTCAGCCCTA 3459
 |||||

QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
 |||||
 Db 3460 CAGCGGTACAGTACAGGATCCACATACCTCTGCTGCCCGCCAGAGCTGATGGTACGTGCT 3519
 |||||

QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
 |||||
 Db 3520 CCCCTGGCTGACAGCCCGCCAGCGAGTATGTGAACCCAGCAGAGGTTCGGCTCAGTCT 3579
 |||||

QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
 |||||
 Db 3580 CCCTTGACCCAGAGGCTCTCCGCTCCCATCCGACCTGCTGGTGTACTCTAGAAAGA 3639
 |||||

QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
 |||||
 Db 3640 CCCAAGACTCTCTCTCTGGGAAAATGGGGTGTCAAGACGTTTTCCTTTGGGGGT 3699
 |||||

QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
 |||||
 Db 3700 GCTGTGGAGAACCTGAATATCTTAGCACCCAGCAGCAGGCTGCTCTCAGCCCCACCT 3759
 |||||

QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
 |||||
 Db 3760 TCTCTGCTTCAGCCAGGCTTTGACAACCTCTATTACTGGGACCAACTCATCGGAG 3819
 |||||

QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
 |||||
 Db 3820 CAGGCTCTCCACCAAGTACTTTGAAGGGGACCCCACTGCAGAGAACCTTGAGTACCTA 3879
 |||||

QY 261 GlyLeuAspValProVal 266
 |||||
 Db 3880 GGCCTGATGTCAGTA 3897
 |||||

RESULT 13
 CD516283
 LOCUS CD516283 852 bp mRNA linear EST 06-JUN-2003
 DEFINITION AGENCOURT_14364504 NIH_MGC_181 Homo sapiens CDNA clone
 IMAGE:30408904 5', mRNA sequence.
 ACCESSION CD516283
 VERSION CD516283.1 GI:31448001
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 852)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.

TITLE
JOURNAL
COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM492 row: j column: 17
High quality sequence stop: 662.
Location/Qualifiers
1. .932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30408904"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_181"
/note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."
BASE COUNT 159 a 288 c 247 g 156 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 2.42e-63 Length: 852
Score: 1165.50 Matches: 237
Percent Similarity: 90.49% Conservative: 1
Best Local Similarity: 90.11% Mismatches: 8
Query Match: 80.38% Indels: 17
Gaps: 3

SEQ4 (1-266) x CD516283 (1-852)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 95 CAGAAAGGAGCTGGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCT 154
|||||

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 155 GAGGACCATGACATGGGGGACCTGGTGTGCTGAGGAGTATCTGGTACCCACGACGGC 214
|||||

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
|||||
Db 215 TTCTCTCTGTCAGACCTGCCCCGGCGCTGGGGCATGGTCCACACAGCAGCCGCGAGC 274
|||||

Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
|||||
Db 275 TCATCTACCAAGAGTGGCGTGGGGACCTGACACTAGGGGTGGAGCCCTCTGAAGAGGAG 334
|||||

Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
|||||
Db 335 GCCCCAGGTCTCACTGGCACCCCTCCAGAGGGGCTGGCTCCGATGATTTGATGGTGAC 394
|||||

Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
|||||
Db 395 CTGGGAATGGGGGAGCAAGGGGCTCAAAAGCCCTCCACACATGACCCCGCCCTCTA 454
|||||

Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
|||||
Db 455 CAGCGGTACAGTAGGAGCCACACAGTACCCCTGCGCCCTCTGAGACTGATGGCTAGTTGCC 514
|||||

Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
|||||

Db 515 CCCCTGACCTGCAGCCCCCAGCCCTGAATATGTGAACCCAGCCAGATGTTGCGCCCCAGCCC 574
|||||

Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
|||||

Db 575 CCTTGGCCCCGAGAGGGCCCTCTGCCTGCTGCCGAGCTGCTGGTGGCACTCTCGAAGG 634
|||||

Qy 181 ProlysThrLeuSerPro-GlyLysAsnGlyValValLysAspValPheAlaPheGlyG1 200
|||||

Db 635 GCCAAGACTCTCTCCCCAGGGAAGAATGGGTGCTCAAGACCGCTTTTGGCTTTGGGG 694
|||||

Qy 200 YAlaValGluAsnProGluTyrLeuThr-ProGlnGlyGlyAla-AlaProGlnProHis 219
|||||

Db 695 TGGCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTTGCCTCAGCCCCAC 754
|||||

Qy 220 -ProProProAla-PheSerProAla---PheAspAsnLeuTyrTyrTrpAspGlnAsp 238
|||||

Db 755 CCCTCTCTCTGCCNTTTCAGCCCGAGCCCTTTCGACAAC-----C 793
|||||

Qy 238 roPro-----GluArgGlyAlaProProSerThrPheLysGlyThrProThr 253
|||||

Db 794 CTCCTATTAACTGGGGAACAGGAGCCCGACCCCAAGAGCGGGGGGCTTCCAAAC 850
|||||

RESULT 14
CA487981

LOCUS CA487981 932 bp mRNA linear EST 14-NOV-2002

DEFINITION AGENCOURT_10808060 MAPcL Homo sapiens cDNA clone IMAGE:6719711 5',
mRNA sequence.

ACCESSION CA487981

VERSION CA487981.1 GI:24948694

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 932)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14278 row: a column: 23
High quality sequence stop: 567.
Location/Qualifiers
1. .932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6719711"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HTERT-HME1
, LNCap"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dr. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."
BASE COUNT 173 a 328 c 268 g 162 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1.67e-61 Length: 932

```

Score: 1136.50 Matches: 229
Percent Similarity: 86.03% Conservative: 5
Best Local Similarity: 84.19% Mismatches: 20
Query Match: 78.38% Indels: 18
DB: 14 Gaps: 5

SEQ4 (1-266) x CA487981 (1-932)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 67 CAGAATGAGGACTTGGCCCGCCAGCCAGTCCCTGGACAGCACCTTCTACCGCTCAGTCTG 126
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
DB 127 GAGGACGATCAGATGGGGACCTGTGTGATGCTGAGGAGTATCTGTATCCCGAGCAGGC 186
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
DB 187 TTCTTCTGTCAGACCTGCCCGGCGCTGGGGCATGTTCCACACAGGCACCCGAGC 246
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB 247 TCATCTACAGGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAGAGGAG 306
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB 307 GCCCCAGGCTCTCACTGGCACCTCCGAAGGGGCTGGCTCCGATGATTTGATGTGAC 366
QY 101 LeuGlyMetClyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
DB 367 CTGGGAATGGGGCGACCAAGGGCTGCAAGGCTCCCCACACATGACCCAGCCCTCTA 426
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
DB 427 CACGGTACAGTACGAGACCCACAGTACCCCTGCTGAGACTGATGCTGCTGCTG 486
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
DB 487 CCCTGACCTGCAGCCCGCCAGCTGAATATGTAACCCAGCCAGATGTTCCGCCCGCCAG 546
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
DB 547 CTTCCGCCCGAGAGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPhe-GlyG 200
DB 607 CCCAACACTCTCTCCCGAGGAGAGATGGGTGCTCAAGACGTTTTTCCCTTTGGGGG 666
QY 200 yAlaValGluAsnProGluTyrLeu-ThrProGlnGly-GlyAlaAlaProGlnProHis 219
DB 667 TGCCGGGAGAACCCCGGACTTGGACCCCGCCAGGGAAGAGCTGGCCCTCAGCCCCAC 726
QY 220 Pro-----ProProAlaPheSerProAlaPheAspAsnLeuTyrTyr 233
DB 727 CTTCTCTTCTGGCTTAGCCCGCCAAACCTTCTAACACCT-----TCTATTTACTGG 777
QY 234 TrpAspGlnAspProGlu---ArgGlyAla-----ProProSerThrPhe----- 248
DB 778 GGACCCCGGGACCCCGCCACAGCGGGGGGGGCGCTCCACCCCGCCACATTTTATGA 837
QY 249 ---LysGlyThrProThrAlaGluAsnPro 257
DB 838 GGGGACCCCTTACCGGCCCGAGATAAACCC 867

RESULT 15
AUI23871
LOCUS AUI23871 NT2RM2 Homo sapiens cDNA clone NT2RM2001211 5', mRNA
DEFINITION sequence.
ACCESSION AUI23871
VERSION AUI23871.1 GI:10948587
KEYWORDS EST.
SOURCE Homo sapiens (human)

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 685)
AUTHORS Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
Y., Sugano,S., Isogai,T.)
JOURNAL Unpublished
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Location/Qualifiers
source
1..685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM2001211"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RM2"
/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
precursor cells"
BASE COUNT 137 a 224 c 198 g 123 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 4..86e-61 Length: 685
Score: 1127.00 Matches: 213
Percent Similarity: 96.83% Conservative: 1
Best Local Similarity: 96.38% Mismatches: 6
Query Match: 77.72% Indels: 1
DB: 9 Gaps: 0
SEQ4 (1-266) x AUI23871 (1-685)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 11 AAGAATGAGGACTTGGGCCCGCCAGCTCCCTTGGACAGCACCTTCTACCGCTCAGTCTG 70
QY 21 GluAspAspMetClyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
DB 71 GAGGACGATGATGAGGGGACCTGGTGGATGCTGAGGAGTATCTGTGTACCCAGCAGGC 130
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
DB 131 TTCTTCTGTCAGACCTGCCCGGCGCTGGGGCATGTTCCACACAGGCACCCGAGC 190
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB 191 TCATCTACAGGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 250
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB 251 GCCCCAGGCTCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGATTTGATGTGAC 310
QY 101 LeuGlyMetClyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
DB 311 CTGGGAATGGGGCGACCAAGGGCTGCAAGGCTCCCGCCACACATGACCCCGCCCTCTA 370
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
DB 371 CAGCGGTACAGTACGAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTACGTTGCC 430

```

QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
 DB 431 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACCCAGCCAGATGTTGGCCCCCAGCCC 490
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
 DB 491 CCTTCGCCCGAGAGGGCCCTGCTGCTGCTGCCGACCTGCTGGTGCCACTCTGAAAGG 550
 QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
 DB 551 GCCAAGACTCTCCCCAGGAGAAAGGGGTCTCAAAAGACGTTTTCCTTTGGGGGT 610
 QY 201 AlaValGluAsnProGluTyr-LeuThrProGlnGlyAlaAlaProGlnProHisPr 220
 DB 611 GCCGTGGAGAACCCCGAGTACTTTGCACACCCCGAGGAGAACTGCCCCCTTAAGCCACC 670
 QY 220 O 220
 DB 671 C 671

Search completed: October 16, 2003, 17:04:49
 Job time : 2095.84 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 11:08:07 ; Search time 274.965 Seconds
(without alignments)
2540.503 Million cell updates/sec

Title: SEQ4

Perfect score: 1450

Sequence: 1 QNEDIGPASPLDSTFYRSL.....TFKGTPTAENPEYLGLDVPV 266

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-DB=PublishedApplications_NA_QFMT-fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=HOLLERAN480 @CGN_1.1.397 @runat.15102003.131915_20662
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PublishedApplications_NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	ID	Description
1	1450	100.0	1713	12	US-10-378-393-14	Sequence 14, Appl

2	1450	100.0	1755	10	US-09-930-125-6	Sequence 6, Appli
3	1450	100.0	1767	10	US-09-930-125-4	Sequence 4, Appli
4	1450	100.0	1773	10	US-09-930-125-7	Sequence 7, Appli
5	1450	100.0	2411	12	US-10-378-393-10	Sequence 10, Appl
6	1450	100.0	3768	10	US-09-854-356-9	Sequence 9, Appli
7	1450	100.0	3768	10	US-09-930-125-1	Sequence 1, Appli
8	1450	100.0	3768	12	US-10-313-644-1	Sequence 5, Appli
9	1450	100.0	4473	11	US-09-441-411-5	Sequence 81, Appl
10	1450	100.0	4473	12	US-10-101-510-81	Sequence 32, Appl
11	1450	100.0	4473	14	US-10-146-473-32	Sequence 44, Appl
12	1450	100.0	4473	14	US-10-207-655-44	Sequence 70, Appl
13	1450	100.0	4606	12	US-09-971-392-70	Sequence 5, Appli
14	1442	99.4	1806	10	US-09-930-125-5	Sequence 8, Appli
15	1442	99.4	3765	12	US-10-207-498-5	Sequence 2, Appli
16	1442	99.4	3768	9	US-09-811-123-8	Sequence 11, Appl
17	1442	99.4	3768	9	US-09-811-115-2	Sequence 119, App
18	1442	99.4	4530	10	US-09-877-177-11	Sequence 124, App
19	1442	99.4	4530	12	US-10-007-926A-119	Sequence 1, Appli
20	1442	99.4	4530	12	US-10-101-510-124	Sequence 125, App
21	1442	99.4	4530	12	US-10-338-730-1	Sequence 1, Appli
22	1442	99.4	4530	14	US-10-177-293-125	Sequence 12, App
23	1442	99.4	4543	10	US-09-769-508-1	Sequence 1, Appli
24	1442	99.4	4642	14	US-10-198-846-10896	Sequence 10896, A
25	1442	99.4	9274	9	US-09-811-123-7	Sequence 7, Appli
26	1442	99.4	9274	9	US-09-811-115-1	Sequence 1, Appli
27	1208	83.3	3955	10	US-09-870-759-117	Sequence 117, App
28	1208	83.3	3955	10	US-09-854-356-10	Sequence 10, Appl
29	1208	83.3	3955	12	US-09-751-708A-117	Sequence 117, App
30	1205	83.1	3771	10	US-09-854-356-11	Sequence 11, Appl
31	1183	81.6	1665	9	US-09-821-883-6	Sequence 6, Appli
32	1183	81.6	2070	9	US-09-821-883-7	Sequence 7, Appli
33	1177	81.2	651	9	US-09-821-883-26	Sequence 26, Appl
34	1177	81.2	1191	9	US-09-821-883-28	Sequence 28, Appl
35	1177	81.2	1692	9	US-09-821-883-8	Sequence 8, Appli
36	1177	81.2	2091	9	US-09-821-883-9	Sequence 9, Appli
37	1177	81.2	3537	9	US-09-821-883-30	Sequence 30, Appl
38	1041.5	71.8	1115	14	US-10-102-806-165	Sequence 165, App
39	790.5	54.5	507	11	US-09-911-904-116	Sequence 116, App
40	674	46.5	452	11	US-09-918-995-32765	Sequence 32765, A
41	647.5	44.7	1046	9	US-09-925-301-390	Sequence 390, App
42	555	38.3	455	11	US-09-918-995-14628	Sequence 14628, A
43	479.5	33.1	343	12	US-10-007-926A-118	Sequence 118, App
44	479.5	33.1	343	12	US-10-007-926A-444	Sequence 444, App
45	435	30.0	519	9	US-09-925-301-642	Sequence 642, App

ALIGNMENTS

RESULT 1
US-10-378-393-14
; Sequence 14, Application US/10378393
; Publication No. US20030182668A1
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVAT
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-14

```

Alignment Scores:
Pred. No.: 5,75e-118 Length: 1713
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ4 (1-266) x US-10-378-393-14 (1-1713)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 913 CAGAATGAGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTACCGCTCACTGCTG 972
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 973 GAGGACGATCAGATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGAGGGC 1032
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 1033 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGTCCACACAGCACCGCAGC 1092
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 1093 TCATCTACAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 1152
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1153 GCCCGCAGGTCTCCAGTGGCACCTCCGAAGGGCGTGGCGCATGTTCGATGATTTGATGGTAC 1212
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 1213 CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCCCACACATGACCCCGCCCTCTA 1272
Qy 121 GluArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 1273 CAGCGGTACAGTGGAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTACGTGGC 1332
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 1333 CCCTTGACCTTGACCCCGCCAGCTGAATATGTGACCCAGCCAGATGTTGGCCCGCCAGCC 1392
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 1393 CCTTCGCCCGCAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGCACCTCTGAAAGG 1452
Qy 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 1453 CCCAAGACTCTCTCCCGAGGAGAAATGGGGTCGTCAAGACAGTCTTGGCTTTGGGGGT 1512
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 1513 GCCGTGGAGAACCCGAGTACTTGCACACCCCGAGGAGACTGCCCTCTAGCCCGCCACCT 1572
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 1573 CCTCCTGCCCTCAGCCCGAGCTTCGACAACCTCTATTACTGGGACAGGACCCAGAG 1632
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 1633 CGGGGGGCTCCACCCAGACCTTCAAGAGGACACCTACGGCAGAGAACCCAGAGTACCTG 1692
Qy 261 GlyLeuAspValProVal 266
Db 1693 GGTCTGGACGTGCCACTG 1710

RESULT 2
US-09-930-125-6
; Sequence 6, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-930-125-6

Alignment Scores:
Pred. No.: 5,9e-118 Length: 1755
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ4 (1-266) x US-09-930-125-6 (1-1755)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 946 CAGAATGAGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTACCGCTCACTGCTG 1005
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 1006 GAGGACGATCAGATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGAGGGC 1065
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGTTCACACAGCACCGCAGC 1125
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 1126 TCATCTACAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 1185
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1186 GCCCGCAGGTCTCCACTGGCACCTCCGAAGGGCGTGGCTCCCATGATGATTTGATGGTAC 1245
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 1246 CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCCCACACATGACCCCGCCCTCTA 1305
Qy 121 GluArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 1306 CAGCGGTACAGTGGAGACCCAGTACCCCTGCCCTCTGAGACTGATGGCTACGTGGCC 1365
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 1366 CCCTGACCTGACAGCCCGCCAGCTGAATATGTGACAGCCAGCATGTTCCGGCCCGCCAGCC 1425
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 1426 CCTTCGCCCGCAGAGGGCCCTCTGCCCTGCTGCCGACCTGCTGGTGCACCTCTGGAAGG 1485
Qy 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 1486 CCCAAGACTCTCTCCCGAGGAGAAATGGGGTCGTCAAGACAGTCTTGGCTTTGGGGGT 1545
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 1546 GCCGTGGAGAACCCGAGTACTTGCACACCCCGAGGAGACTGCCCTCTAGCCCGCCACCT 1605
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240

```

```
Db 1606 CCTCTGCTTTCAGCCAGCTTCGACAACTCTATTACTGGGACGAGCCACCAGAG 1665
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 1666 CGGGGGCTCCACCCAGACCTTCAAAGGACACCTACGGCAGAGAAACCCAGAGTACCTG 1725
QY 261 GlyLeuAspValProVal 266
Db 1726 GGTCTGGAGCTGCCAGTG 1743

RESULT 3
US-09-930-125-4
; Sequence 4, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedwick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-4

Alignment Scores:
Pred. No.: 5,94e-118 Length: 1767
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ4 (1-266) x US-09-930-125-4 (1-1767)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 946 CAGAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 1006 GAGGACGATGACATGGGGACCTGTGTGATGCTGAGGAGTATCTGTGTACCCAGCAGGCG 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 1066 TTCTTCTGCCAGACCTTCCCGGGCGCTGGGGCATGGTCCACCCAGCAGCCAGC 1125
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 80
Db 1126 TCATCTACCAAGGAGTGGGGTGGGACCTGACACCTAGGCTGGAGCCCTCTCAAGAGGAG 1185
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1186 GCCCCCAGGTCTCCACTGGCACCCCTCCGAAGGGGCTGGCTCCGATGTATTGTATGGTGAC 1245
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisHisAspProSerProLeu 120
Db 1246 CTGGGAATGGGGGACCCAGAGGGGTGCAAAGCCCTCCCCACACATGACCCACCCCTCTA 1305
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 1306 CAGCGGTACAGTGGAGGACCCACAGTACCCCTGCCCTCTGAGACTGTAGGCTACGCTTGCC 1365

US-09-930-125-7
; Sequence 7, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedwick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-7

Alignment Scores:
Pred. No.: 5,96e-118 Length: 1773
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ4 (1-266) x US-09-930-125-7 (1-1773)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 970 CAGAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1029
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 1030 GAGGACGATGACATGGGGACCTGTGTGATGCTGAGGAGTATCTGTGTACCCAGCAGGCG 1089
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
```

```
Db 1090 TTCCTCTGTCAGACCTGCCCCGGGGCTGGGGCATGGTCCACACAGCAGCCGCGAGC 1149
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 1150 TCATCTACAGAGTGGGGTGGGGAGCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 1209
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1210 GCCCCAGGTCTCACCTGGCACCTCCGAAGGGCTGGCTCCGATGATTATGGTGAC 1269
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 1270 CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCACACATGATGATGATGATGATGAT 1329
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 1330 CAGCGGTACAGTGGAGACCCACAGTACCCCTGCCCCCTGAGACTGATGGCTAGGTTGCC 1389
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 1390 CCCCTGACCTGCAGCCCCCAGCCCTGATATGTGAACCCAGCCAGATGTTGCGCCCCAGCCC 1449
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 1450 CCTTCGCCCCGAGAGGGCCCTCTGCCTGCTGCCGACCTGCTGGTGCCACTCTGGAAGG 1509
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 1510 CCCAAGACTCTCTCCCCAGGGAAGATGGGGTCTGCAAGACGTTTTTGGCTTTGGGGGT 1569
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 1570 GCCGTGAGAACCCGAGTACTTGACACCCAGGGAGGAGCTGCCCTCAGCCCCACCT 1629
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 1630 CCTCCTGCCCTCAGCCAGCCCTTGCAACACTCTATTACTGGGACCCAGGACCCACCCAGAG 1689
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 1690 CGGGGGGCTCCACCCAGCCTTCAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 1749
Qy 261 GlyLeuAspValProVal 266
Db 1750 GGTCTGACGTGCCAGTG 1767
```

RESULT 5

```
US-10-378-393-10
; Sequence 10, Application US/10378393
; Publication No. US2003018268A1
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
; TITLE OF INVENTION: TYROSINE KINASE RECEPTORS
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; PRIOR FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-10
```

Alignment Scores:

```
Pred. No.: 8,12e-118 Length: 2411
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ4 (1-266) x US-10-378-393-10 (1-2411)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 1611 CAGAATGAGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1670
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 1671 GAGGACATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCACGAGGC 1730
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 1731 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGTGTCCACACAGCAGCAGCCGAGC 1790
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 1791 TCATCTACCAAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 1850
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1851 GCCCCAGAGTCTCCACTGGGACCCCTCCGAAGGGCTGGCTCCGATGATTATGGTGAC 1910
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 1911 CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCACACATGATGATGATGATGATGAT 1970
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 1971 CAGCGGTACAGTGGAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTAGTTGCC 2030
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 2031 CCCCTGACCTGCAGCCCCCAGCCCTGAATATGTGAACCCAGCCAGATGTTGCGCCCCAGCCC 2090
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 2091 CCTTCGCCCCGAGAGGGCCCTCTGCCCTGCTGCCGACCTGCTGGTGCCACTCTGGAAGG 2150
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 2151 CCCAAGACTCTCTCCCCAGGGAAGATGGGGTCTGCAAGACGTTTTTGGCTTTGGGGGT 2210
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 2211 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGAGCTGCCCTCAGCCCCACCCCT 2270
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 2271 CCTCCTGCCCTTCAGCCAGCCCTTCGACAACTCTATTACTGGGACCCAGGACCCACCCAGAG 2330
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 2331 CGGGGGGCTCCACCCAGCACCTTCAAGGGGACACCTACGGCAGAGAACCCAGAGTACCTG 2390
Qy 261 GlyLeuAspValProVal 266
Db 2391 GGTCTGACGTGCCAGTG 2408

RESULT 6
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
```

```

; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: Intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.: 1.28e-117 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ4 (1-266) x US-09-854-356-9 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGGACTTGGGCGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTCTGTCAGACCTCGCCCGGGCGCTGGGGCATGTGTCACACAGCAGCCGCGAGC 3147

Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3148 TCATCTACGAGAGTGGCGGTGGGACCTGACATAGGCTGGAGCCCTCTGAGAGAGG 3207

Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCGCCAGGCTCTCCATGGCACCTCCGAGGGGCTGGCTCGATGATTTGATGGTGAC 3267

Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGAATGAGGCGGAGCCAGGCGGTGCAAGGCTGCAAGGCTCCACACATGACCCAGCCCTCTA 3327

Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGGTACAGTGAGGACCCACAGTACCCCTGCTGAGCTGATGGCTACGTGGC 3387

Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160

```

```

Db 3388 CCCCTGACCTGACGCCCCAGCCTGAATATGTGAACAGCCAGATGTTTGGCCCCAGGCC 3447
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3448 CCTCGCCCCAGAGGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3507
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3508 CCCAAGACTCTCTCCCGAGGAAGAAATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3567
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3568 GCGTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGGTGCTGCTGCTGCTGCTGCTGCTG 3627
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3628 CCTCTGCTTTCAGCCCGAGCCTTCGACAACTCTATTACTGGGACGAGGACCCAGCAGAG 3687
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3688 CGGGGGCTCCAGCCAGCACCTTCAAAGGGACACCTACGGCAGAGACCCAGAGTACCTG 3747
Qy 261 GlyLeuAspValProVal 266
Db 3748 GGTCTGGAGTGGCCAGTG 3765

RESULT 7
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930.125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-09-930-125-1

Alignment Scores:
Pred. No.: 1.28e-117 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ4 (1-266) x US-09-930-125-1 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGGACTTGGGCGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTCTGTCAGACCTCGCCCGGGCGCTGGGGCATGTGTCACACAGCAGCCGCGAGC 3147

Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3148 TCATCTACGAGAGTGGCGGTGGGACCTGACATAGGCTGGAGCCCTCTGAGAGAGG 3207

Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCGCCAGGCTCTCCATGGCACCTCCGAGGGGCTGGCTCGATGATTTGATGGTGAC 3267

Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGAATGAGGCGGAGCCAGGCGGTGCAAGGCTGCAAGGCTCCACACATGACCCAGCCCTCTA 3327

Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGGTACAGTGAGGACCCACAGTACCCCTGCTGAGCTGATGGCTACGTGGC 3387

Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160

```

```

Qy 41 PhePheCysProaspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTCTGTGCAGACCCCTGCCCGGGGCTGGGGGATGTTCCACACAGCAGCCGCGC 3147
Qy 61 SerSerThrArgSerGlyGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 80
Db 3148 TCATCTACAGAGGAGTGGCGTGGGACCTCACACTAGGGGTGGAGCCCTCTGAAGAGGAG 3207
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCCCCAGGCTCCACTGGCACCCCTCGAAGGGGCTGGCTCCCATGTATTTGATGTGAC 3267
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuInSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGATGGGGGAGCAAGGGGCTGCAAAAGCTCCACACATGACCCAGCCCTCTA 3327
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGGTACAGTGAAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTACGTGGC 3387
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3388 CCCTGACCTGCAGCCGCCAGCTGAATATGTGAACCCAGCCAGATGTTCCGCCCCAGCCC 3447
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3448 CCTTCGCGCCGAGAGGGCCCTCTGCCTGCTGCGCCGACCTGCTGGTGCCACTCTGAAAGG 3507
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3508 CCCAAGACTCTCTCCCGAGGAGATGGGTCTCAAGACCGTTTGGCTTTGGGGGT 3567
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3568 GCGGTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGCTGCCCTCAGCCCCACCCCT 3627
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3628 CTTCTCGCTTCAGCCCGAGCTTCGACAACTCTATTTACTGGGACCCAGGACCCAGAG 3687
Qy 241 ArgGlyAlaProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3748 GGTCTGGACGTGCCAGTG 3765

```

RESULT 8

```

US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313,644
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-313-644-1

```

Alignment Scores:

```

Pred. No.: 1,28e-117 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ4 (1-266) x US-10-313-644-1 (1-3768)

Qy 1 GlnAspGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGGACTTGGGCCCGAGCCAGTCCCTTGGACAGCACCTTCTACCCCTCAGTCTG 3027
Qy 21 GlnAspAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProaspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTCTGTGCAGACCCCTGCCCGGGGCTGGGGGATGTTCCACACAGCAGCCGCGC 3147
Qy 61 SerSerThrArgSerGlyGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3148 TCATCTACAGAGGAGTGGCGTGGGACCTGACACTAGGGGTGGAGCCCTCTGAAGAGGAG 3207
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCCCCAGGCTCCACTGGCACCCCTCGAAGGGGCTGGCTCCCATGTATTTGATGTGAC 3267
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGATGGGGGAGCAAGGGGCTGCAAAAGCTCCACACATGACCCAGCCCTCTA 3327
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGGTACAGTGAAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTACGTGGC 3387
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3388 CCCTGACCTGCAGCCGCCAGCTGAATATGTGAACCCAGCCAGATGTTCCGCCCCAGCCC 3447
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3448 CCTTCGCGCCGAGAGGGCCCTCTGCCTGCTGCGCCGACCTGCTGGTGCCACTCTGAAAGG 3507
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3508 CCCAAGACTCTCTCCCGAGGAGATGGGTCTCAAGACCGTTTGGCTTTGGGGGT 3567
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3568 GCGGTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGCTGCCCTCAGCCCCACCCCT 3627
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3628 CTTCTCGCTTCAGCCCGAGCTTCGACAACTCTATTTACTGGGACCCAGGACCCAGAG 3687
Qy 241 ArgGlyAlaProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3688 CGGGGGGCTTCACCCAGACCTTCAAAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3747
Qy 261 GlyLeuAspValProVal 266
Db 3748 GGTCTGGACGTGCCAGTG 3765

```

RESULT 9

```

US-09-441-411-5
; Sequence 5, Application US/09441411
; Publication No. US2003008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd

```

; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-5

Alignment Scores:
Pred. No.: 1,52e-117 Length: 4473
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

SEQ4 (1-266) x US-09-441-411-5 (1-4473)

```
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGAATGAGGACTTGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGly 40
Db 3202 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 3262 TTCCTCTCCAGACCTGCCCCGGCGCTGGGGGCGATGGTCCACACAGCAGCAGCGCAGC 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3322 TCATCTACGAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3382 CCCCCAGGCTCTCCACTGGCACCTCCCGAGGGGCTGGCTCGGATGATTTGATGGTGAC 3441
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3442 CTGGGAATGGGGGACCCAGGGGCTGCAAGCCCTCCACACATGACCCCGCCCTCTA 3501
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3502 CAGCGGTACAGTACAGGACCCACAGTACCCCTGCGCTCTGAGACTGATGGCTACGTTGCC 3561
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3562 CCCCTGACTCGAGCCCCCAGCCTGAATATGTGAACACGACAGATGTTGGGCCCGCAGCCC 3621
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3622 CTTTGGCCCGAGAGGGCCCTCGCTGCTGCCGACCTGCTGGTGCCACTCTGGAAAGG 3681
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3682 CCCAAGACTCTCTCCCGAGGAAGAAATGGGTGCTGCTCAAAAGACGTTTTTGGCCTTGGGGGT 3741
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 3742 CCGGTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGGCTGCCCCCTCAGCCCCACCCCT 3801
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3802 CCTCTGCTTCAGCCCCAGCCTTCGACAACTCTATTACTGGGACGAGCAGCAGCAGCAG 3861
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
|||||
```

```
Db 3862 CGGGGGGCTCCACCCAGCACCTTCAAAGGAGCACCTACGGCAGAGAACCCAGAGTACCTG 3921
QY 261 GlyLeuAspValProVal 266
|||||
Db 3922 GGTGGACAGTGGCCAGTG 3939
RESULT 10
US-10-101-510-81
; Sequence 81, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-81
Alignment Scores:
Pred. No.: 1,52e-117 Length: 4473
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
```

SEQ4 (1-266) x US-10-101-510-81 (1-4473)

```
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGAATGAGGACTTGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 3262 TTCCTCTCCAGACCTGCCCCGGCGCTGGGGGCGATGGTCCACACAGCAGCAGCGCAGC 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3322 TCATCTACGAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3382 CCCCCAGGCTCTCCACTGGCACCTCCCGAGGGGCTGGCTCGGATGATTTGATGGTGAC 3441
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3442 CTGGGAATGGGGGAGCCCAAGGGGCTGCAAAAGCCCTCCACACATGACCCCGCCCTCTA 3501
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3502 CAGCGGTACAGTACGAGCCCGACAGTACCCCTGCGCTCTGAGACTGATGGCTACGTTGCC 3561
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3562 CCCCTGACTCGAGCCCCCAGCCTGAATATGTGAACACGACAGATGTTGGGCCCGCAGCCC 3621
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3622 CTTTGGCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGCCACTCTGGAAAGG 3681
```


Db	3382	GC	CCCCAGGCTCCACTGGCACCCTCCGAAGGGGCTGGCTCCGATGATTATTTGATGGTGAC	3441
Qy	101	Leu	GlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu	120
Db	3442	CT	GGAAATGGGGCAGCCAAAGGGCTGCAAGGCTCCCCACATGACCCAGCCCTCTA	3501
Qy	121	Gln	ArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla	140
Db	3502	CAG	CGTACAGTGAGGACCCACAGTACCCTGCCCTCTGAGACTGATGGCTACGTTGCC	3561
Qy	141	Pro	LeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro	160
Db	3562	CC	CTGACCTGCAGCCCCCAGCTGAAATATGTGAACACGACAGATGTTCGGCCCCAGCC	3621
Qy	161	Pro	SerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg	180
Db	3622	CC	TTCGGCCCCGAGGGCCCTCTGCCCTGCTGCCGACCTGCTGGTGGCACTCTGGAAGG	3681
Qy	181	Pro	LysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly	200
Db	3682	CC	CAAGACTCTCCCCAGGGAAGATGGTGGTCAACACGCTTTTGGCTTTGGGGGT	3741
Qy	201	Ala	ValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaProGlnProHisPro	220
Db	3742	GC	GTGAGAACCCCGAGTACTTGACACCCAGGAGAGCTGCCCTCAGCCCCACCC	3801
Qy	221	Pro	ProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu	240
Db	3802	CT	TCTGCCCTTCAGCCCGAGCTTCGCAACCTCTATTACTGGGACACAGACCACAGAG	3861
Qy	241	Arg	GlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu	260
Db	3862	CG	GGGGGCTCCACCCAGCACCTTCAAGGGACACCTACGGCAGAGACCCAGAGTACCTG	3921
Qy	261	Gly	LeuAspValProVal 266	
Db	3922	GG	TCTGGACGTGCCAGTG 3939	
RESULT 12				
US-10-207-655-44				
; Sequence 44, Application US/10207655				
; Publication No. US20030118592A1				
; GENERAL INFORMATION:				
; APPLICANT: Ledbetter, Jeffrey A.				
; APPLICANT: Hayden-Ledbetter, Martha S.				
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS				
; FILE REFERENCE: 390069.401C1				
; CURRENT APPLICATION NUMBER: US/10/207.655				
; CURRENT FILING DATE: 2002-07-25				
; NUMBER OF SEQ ID NOS: 426				
; SOFTWARE: PatentIn version 3.0				
; SEQ ID NO 44				
; LENGTH: 4473				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
US-10-207-655-44				
Alignment Scores:				
Pred. No.:		1.52e-117	Length:	4473
Score:		1450.00	Matches:	266
Percent Similarity:		100.00%	Conservative:	0
Best Local Similarity:		100.00%	Mismatches:	0
Query Match:		100.00%	Indels:	0
DB:		14	Gaps:	0
SEQ4 (1-266) x US-10-207-655-44 (1-4473)				
Qy	1	Gln	AsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu	20
Db	3142	CAG	AATGAGGACTTGGGCCCCAGCCAGTCCCTTGACAGCACCTTCTACCCGCTCACTGCTG	3201
Qy	21	Glu	AspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly	40

```
Db 3202 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3261
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60
Db 3262 TTCCTCTCTCAGACCCCTGCCCCGGGCGCTGGGGCATGGTCCACACAGCAGCCGCGAGC 3321
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3322 TCATCTACGAGGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAGAGGAG 3381
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3382 GCCCCAGGCTCTCCACTGGCACCTCCGAAAGGGCTGGCTCCGATGTATTTGATGTTGAC 3441
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3442 CTGGAAATGGGGGAGCCCAAGGGCTGCAAAAGCCTCCCCACATGACCCAGCCCTCTA 3501
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3502 CAGCGGTACAGTGGAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTAGTTGCC 3561
Qy 141 ProLeuThrCysSerProGlnProGluTyrValLysGlnProAspValArgProGlnPro 160
Db 3562 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACACAGCCAGATGTTCCGGCCAGGCC 3621
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3622 CCTTCGCCCCAGAGGGCCCTCTGCTCTGCCGACCTGCTGGTGCCACCTCTCGAAAGG 3681
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValLysAspValPheAlaPheGlyGly 200
Db 3682 CCCAAGACTCTCTCCCGAGGAGAAATGGGTCGTCGCAAGACGTTTTTGGCGGT 3741
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3742 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGTGCCTCTACGCCCCACCC 3801
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3802 CCTCTGCTTCAGCCGAGCCTTCGACAACTCTATTACTGGGACGAGCCACCCAGAG 3861
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3862 CGGGGGGCTCCACCAGCACCTTCAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3921
Qy 261 GlyLeuAspValProVal 266
Db 3922 GGTCTGGACGTGCCAGTG 3939
```

RESULT 13

```
US-09-971-392-70
; Sequence 70, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Pearson, Cecelia I.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971.392
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 70
; LENGTH: 4606
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Template ID: 276948.4
```

```
US-09-971-392-70
Alignment Scores:
Pred. No.: 1.56e-117 Length: 4606
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Db: 12 Gaps: 0

SEQ4 (1-266) x US-09-971-392-70 (1-4606)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3188 CAGAAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACAGACACCTTCTACCGCTCACTGCTG 3247
Qy 21 GluAspAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3248 CAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3307
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 3308 TTCCTCTCTCCAGACCCCTGCCCGGGCGCTGGGGCATGGTCCACACAGGACCCGCGAGC 3367
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3368 TCATCTACGAGGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3427
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3428 GCCCCAGGCTCTCCACTGGCACCTCCGAAAGGGCTGGCTCCGATGTATTTGATGTTGAC 3487
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3488 CTGGAAATGGGGGAGCCCAAGGGCTGCAAAAGCCTCCCCACATGACCCAGCCCTCTA 3547
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3548 CAGCGGTACAGTGGAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTACCTGCC 3607
Qy 141 ProLeuThrCysSerProGlnProGluTyrValLysGlnProAspValArgProGlnPro 160
Db 3608 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACACGAGATGTTTCGGCCCCAGGCC 3667
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3668 CCTTCGCCCCAGAGGGCCCTCTGCTCTGCCGACCTGCTGGTGCCACCTCTGGAAAGG 3727
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValLysAspValPheAlaPheGlyGly 200
Db 3728 CCCAAGACTCTCTCCCGAGGAGAAATGGGTCGTCGCAAGACGTTTTTTCCTTTGGGGGT 3787
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3788 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCTCTCAGCCCCACCT 3847
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3848 CCTCTGCTTCAGCCCCAGCCCTTCGACAACTCTATTACTGGGACGAGACCCAGCAGAG 3907
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3908 CGGGGGGCTCCACCAGCACCTTCAAGGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3967
Qy 261 GlyLeuAspValProVal 266
Db 3968 GGTCTGGACGTGCCAGTG 3985
```

RESULT 14

```
US-09-930-125-5
; Sequence 5, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930.125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-5

Alignment Scores:
Pred. No.:      3,04e-117      Length:      1806
Score:          1442.00      Matches:      265
Percent Similarity: 99.62%      Conservative: 0
Best Local Similarity: 99.62%      Mismatches: 1
Query Match:      99.45%      Indels:      0
DB:              10          Gaps:      0

SEQ4 (1-266) x US-09-930-125-5 (1-1806)

Qy      1  GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db      1003 CAGAATGAGGACTTGGCCCGCAGCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1062

Qy      21  GluAspAspMetGlyAspLeuValAlaGluGlyTyrLeuValProGlnGlnGly 40
Db      1063 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGAGATATCTGGTACCCAGCAGGGC 1122

Qy      41  PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db      1123 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGTGTCACACAGCAGCCGCGCAG 1182

Qy      61  SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db      1183 TCATCTACCAAGAGTGGCGTGGGACCTGACACTAGGGCTGAGGCCCTCTGAAGAGGAG 1242

Qy      81  AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db      1243 GCGCCCGAGGCTCCACTGGGACCCCTCGGAAGGGGCTGGCTCCGATGATTTGATGGTGAC 1302

Qy      101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db      1303 CTGGGAATGGGGCAGCAGCAAGGGCTGCAAAAGCCTCCCCACACATGACCCCGCCCTCTA 1362

Qy      121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db      1363 CAGCGGTACAGTGAAGACCCAGTACCCCTCCCTCTGAGACTGATGGCTACGTTGCC 1422

Qy      141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db      1423 CCCTGACCTGACGCCGCCAGCCTGAATATGTGAACAGGACGATGTTTCGGCCCGCAGCC 1482

Qy      161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db      1483 CCTTGGCCCGCAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGCTGCTGCTGCTGCTGCTG 1542

Qy      181 ProIlystThrLeuSerProGlyLysAsnGlyValValIlystAspValPheAlaPheGlyGly 200
Db      1543 GCCAAGACTCTCTCCCGAGGAGAAATGGGGTGGTCTCAAGACAGCTTTTTCGCTTTGGGGGT 1602

Qy      201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db      1603 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCTCTACGCCCCACCTT 1662

```

```

Qy      221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db      1663 CCTCCTCCCTTCAGCCCGCAGCTTCGACAACCTCTATTACTGGGACGAGGCCACCCAGAG 1722

Qy      241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db      1723 CGGGGGCTCCACCCAGCACCTTCAAAGGAGACCTACGGCAGAGAACCCAGAGTACCTG 1782

Qy      261 GlyLeuAspValProVal 266
Db      1783 GGTCTGGACGTGCCAGTG 1800

RESULT 15
US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; FILE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREGULIN AND HER3
; FILE REFERENCE: 30448.103-US-U1
; CURRENT APPLICATION NUMBER: US/10/207.498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5

Alignment Scores:
Pred. No.:      6.38e-117      Length:      3765
Score:          1442.00      Matches:      265
Percent Similarity: 99.62%      Conservative: 0
Best Local Similarity: 99.62%      Mismatches: 1
Query Match:      99.45%      Indels:      0
DB:              12          Gaps:      0

SEQ4 (1-266) x US-10-207-498-5 (1-3765)

Qy      1  GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db      2968 CAGAATGAGGACTTGGCCCGCAGCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027

Qy      21  GluAspAspMetGlyAspLeuValAlaGluGlyTyrLeuValProGlnGlnGly 40
Db      3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGAGTATCTGGTACCCAGCAGGGC 3087

Qy      41  PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db      3088 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGATGTGTCACACAGACCCCGCAGC 3147

Qy      61  SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db      3148 TCATCTACCAAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3207

Qy      81  AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db      3208 GCGCCCGAGGCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGATTTGATGGTGAC 3267

Qy      101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db      3268 CTGGGAATGGGGGAGCAAGGGGCTGCAAAAGCCTCCCCACACATGACCCCGCCCTCTA 3327

```

QY	121	GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla	140
Db	3328	CAGCGGTACAGTGAGACCCACACAGTACCCCTGCCCTGTGACATGATGGCTACGTTGCC	3387
QY	141	ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro	160
Db	3388	CCCTTGACCTGCAGCCCGCCAGCCTGAATATGTGAACCCAGCCAGATGTTGGGCCCCAGCCC	3447
QY	161	ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrIeuGluArg	180
Db	3448	CTTTCGCCCCGAGAGGGCCCTCTGCTGCTGCCCGACCTGTGGTGCCTCTGGAAAGG	3507
QY	181	ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly	200
Db	3508	GCCAAAGACTCTCTCCCGAGGGAAGATGGGGTCTGTCAAAGACGTTTTTTCCTTTGGGGGT	3567
QY	201	AlaValGluAsnProGluTyrIeuThrProGlnGlyGlyAlaAlaProGlnProHisPro	220
Db	3568	GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCTCAGCCCCACCCCT	3627
QY	221	ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu	240
Db	3628	CCTCTGCTCTCAGCCCGAGCCTTCGACAACTCTATTACTGGGACCGAGCCACCCAGAG	3687
QY	241	ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu	260
Db	3688	CGGGGGGCTCCACCCAGCACCTTCAAGGGGACACCTACGGCAGAGAACCCAGATACCTG	3747
QY	261	GlyLeuAspValProVal	266
Db	3748	GGTCTGGAGTGCCAGTG	3765

Search completed: October 17, 2003, 03:54:36
Job time : 299.965 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 252.648 Seconds
(without alignments)
2842.104 Million cell updates/sec

Title: SEQ4
Perfect score: 1450
Sequence: 1 QNEDLGASPLDSTFYRSL.....TFKGTPTAENPEYLGLDVVPV 266

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/HOLLERAN480/runat_15102003_131912_20526/app_query.fasta_1.4685
-DB=N_Geneseq_19Jun03 -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGN 1.1_1758 -@runat_15102003_131912_20526 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1450	100.0	1755	24	AAD32746	Human CDNA for the
2	1450	100.0	1767	24	AAD32744	Human CDNA for the
3	1450	100.0	1773	24	AAD32747	Human CDNA for the
4	1450	100.0	3768	17	AA740739	HER-2/neu oncogene
5	1450	100.0	3768	20	AAH01912	Human HER-2/neu on
6	1450	100.0	3768	22	AAH23392	Human HER-2/neu pr
7	1450	100.0	3768	24	AAD32743	Human Her-2/neu pr
8	1450	100.0	3768	24	ABA92253	Human Her-2/neu CD
9	1450	100.0	3768	24	ABK10730	Human Her-2/neu DN
10	1450	100.0	4299	14	AAQ46083	Sequence encoding
11	1450	100.0	4472	21	AAA14812	CDNA encoding the
12	1450	100.0	4473	19	ABQ76220	Human tumour anti
13	1450	100.0	4473	20	AAZ31071	HER-2 nucleic acid
14	1450	100.0	4473	24	ABZ34969	Human gene express
15	1450	100.0	4473	24	ABZ34969	Human Her-2 DNA.
16	1442	99.4	1806	24	AAD32745	Human CDNA for the
17	1442	99.4	3678	24	ABK86207	CDNA encoding huma
18	1442	99.4	3768	21	AAA09455	Human heregulin 2
19	1442	99.4	3768	24	ABZ35744	Human ERBB2 polyu
20	1442	99.4	3768	24	ABX09887	Human ERBB2 DNA fr
21	1442	99.4	3768	24	AAD43935	Human HER-2 CDNA.
22	1442	99.4	3768	24	AAD43986	Human Her2 antigen
23	1442	99.4	3768	24	ABV78168	Human ERBB2 DNA SE
24	1442	99.4	3768	24	ABL91709	Human polynucleoti
25	1442	99.4	3768	24	ABK14058	Human HER2 (Erbb2)
26	1442	99.4	4530	16	AA701585	Her-2/neu (ERBB2/c
27	1442	99.4	4530	18	AA71253	Human HER2 gene.
28	1442	99.4	4530	21	AAZ60815	Nucleotide sequenc
29	1442	99.4	4530	22	AAD19731	Human tyrosine kin
30	1442	99.4	4530	24	ABZ35012	Human gene express
31	1442	99.4	4530	24	ABV94128	Breast carcinoma r
32	1442	99.4	4530	24	ABN85585	Human HER2-neu SEQ
33	1442	99.4	4530	24	ABK83918	Human CDNA differe
34	1442	99.4	4530	25	ACC50139	Breast cancer asso
35	1442	99.4	4530	25	ABQ83856	Human Her2/Neu enc
36	1442	99.4	9274	24	AAD43934	HER-2 transgene pl
37	1442	99.4	9274	24	ABK14057	Human HER2 (Erbb2)
38	1208	83.3	3955	16	AA701590	HER2 transgene pla
39	1208	83.3	3955	16	AA701590	Rat neu promoter.
40	1208	83.3	3955	21	AAA89753	Rat HER-2/neu prot
41	1205	83.1	2125	19	AAV21727	Humanised vector p
42	1205	83.1	2763	24	ABA92252	Mouse Her-2/neu ex
43	1205	83.1	2781	24	ABA92253	Her-2/neu extracel
44	1205	83.1	3771	21	AAA89737	Mouse Her-2/neu CD
45	1205	83.1	3771	22	AAH42210	Nucleotide sequenc

ALIGNMENTS

RESULT 1

AAD32746
ID AAD32746 standard; CDNA; 1755 BP.

AC AAD32746;

DT 01-JUL-2002 (first entry)

DE Human CDNA for the clone HICD_native_coding_region.

XX Human; Her-2/Neu protein; immune response; breast cancer;
KW human leukocyte antigen; HLA; vaccine; malignancy; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1755

```

FT      /*tag= a
FT      /product= "Human protein encoded by cDNA for the clone
FT      HICD_native_coding_region"
FT      /trans_except= (pos:1741..1752, aa:Leu-Glu)
FT      /note= "CDS does not include stop codon"
FT      /partial
XX
XX      WO200214503-A2.
XX
XX      21-FEB-2002.
XX
XX      14-AUG-2001; 2001WO-US41733.
XX
XX      14-AUG-2000; 2000US-225152P.
XX      28-SEP-2000; 2000US-236428P.
XX      21-FEB-2001; 2001US-270520P.
XX
XX      (CORI-) CORIXA CORP.
XX
XX      Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
XX      Mcneill PD, Vedvick TS;
XX
XX      WPI; 2002-280758/32.
XX      P-PSDB; RAE20483.
XX
XX      Novel isolated Her-2/Neu polypeptide composition useful for therapy,
XX      prevention and diagnosis of cancer, preferably breast cancer -
XX
XX      Example 5; Page 118-119; 129pp; English.
XX
XX      The invention relates to an isolated Her-2/Neu polypeptide composition
XX      effective for eliciting an immune response. The invention is useful for
XX      eliciting an immune response in a patient, where the patient is human
XX      leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
XX      The composition is useful for the therapy and diagnosis of cancer.
XX      preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
XX      and other compositions for the diagnosis, prevention and treatment of
XX      human malignancies, for stimulating and/or expanding T cells specific for
XX      Her-2/Neu polypeptide and for inhibiting the development of cancer in a
XX      patient. The invention is useful for stimulating a T cell response in a
XX      human patient, as probe or primer for nucleic acid hybridisation, to
XX      selectively form duplex molecules with complementary stretches of the
XX      entire Her-2/Neu gene or gene fragments of interest, to isolate a full
XX      length gene from a suitable library, and to direct expression of a
XX      polypeptide in appropriate host cells. The composition is useful in
XX      prophylactic or therapeutic applications and for the treatment of cancer,
XX      preferably for the immunotherapy of breast cancer and other Her-2/Neu-
XX      associated malignancies. The invention is useful in gene therapy. The
XX      present sequence is human cDNA for the clone HICD_native_coding_region.
XX
XX      Sequence 1755 BP; 376 A; 517 C; 531 G; 331 T; 0 other;

Alignment Scores:
Pred No.:      4.69e-67      Length:      1755
Score:      1450.00      Matches:      266
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      24      Gaps:      0

SEQ4 (1-266) x AAD32746 (1-1755)

QY      1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db      946 CAGAAATGAGGACITGGGCCCCAGCAGTCCCTTGGACAGACCTTCTACCGCTCACTGCTG 1005
QY      21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db      1006 GAGGACCATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGC 1065
QY      41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
Db      1066 TTCTTCTGTCCAGACCTGCCCGGGGCGCTGGGGGGCATGGTCCACCAGACGCGCAGC 1125

```

```

QY      61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db      1126 TCATCTACCAAGAGTGGCGGTGGGACCTGACACTAGGCTGGAGCCCTCTGAAGAGGAG 1185
QY      81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db      1186 GCCCCAGGCTCCACCTGGCACCCCTCCGAAAGGGCTGGCTCGATGATTTGATGGTGAC 1245
QY      101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db      1246 CTGGGAATGGGGGAGCCCAAGGGCTGCAAAAGCCTCCCCACATGAGCCCGAGCCCTCTA 1305
QY      121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db      1306 CAGCGGTACAGTGAGGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTAGCTTCC 1365
QY      141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db      1366 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACCCAGCCAGATGTTGGCCCGCAGCCC 1425
QY      161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db      1426 CCTTCGCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGCCACTCTGGAAGG 1485
QY      181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db      1486 CCCAAGACTCTCTCCCCAGGGAAGATGGGGTCTGCAAAAGACGTTTTTGCCTTTGGGGGT 1545
QY      201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db      1546 GCCGTGGAGAACCCTGAGTACTTGACACCCCGAGGAGGAGTGGCCCTTCAGCCCGCACCT 1605
QY      221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db      1606 CCTCCTCGCTTCAGCCCGACCTTCGACAACCTCTATTACTGGGACCCAGACCCACCCAGAG 1665
QY      241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db      1666 CGGGGGGCTCCACCCAGCACCTTCAAAGGGACACCTACGGCAGAGAGCCAGAGTACCTG 1725
QY      261 GlyLeuAspValProVal 266
Db      1726 GGTCTGGACGTGCCAGTG 1743
RESULT 2
AAD32744
ID      AAD32744 standard; cDNA; 1767 BP.
XX
XX      AAD32744;
XX
XX      01-JUL-2002 (first entry)
XX
XX      Human cDNA for the clone HICD_CT_His_coding_region.
XX
XX      Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
XX      human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
XX
XX      Homo sapiens.
XX
XX      Key      Location/Qualifiers
XX      CDS      1..1764
XX      /tag= a
XX      /product= "Human protein encoded by cDNA for the clone
XX      HICD_CT_His_coding_region"
XX
XX      WO200214503-A2.
XX
XX      21-FEB-2002.
XX
XX      14-AUG-2001; 2001WO-US41733.
XX
XX      14-AUG-2000; 2000US-225152P.
XX
XX      14-AUG-2000; 2000US-225152P.

```


PR 28-SEP-2000; 2000US-236428P.
PR 21-FEB-2001; 2001US-270520P.
XX (CORI-) CORIXA CORP.
XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
PI McNeill PD, Vedvick TS;
XX WPI: 2002-280758/32.
DR P-PSDB; AAE20481.
XX
PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,
PT prevention and diagnosis of cancer, preferably breast cancer
XX
PS Example 5; Page 117-118; 129pp; English.
XX
CC The invention relates to an isolated Her-2/Neu polypeptide composition
CC effective for eliciting an immune response. The invention is useful for
CC eliciting an immune response in a patient, where the patient is human
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
CC The composition is useful for the therapy and diagnosis of cancer,
CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
CC and other compositions for the diagnosis, prevention and treatment of
CC human malignancies, for stimulating and/or expanding T cells specific for
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
CC patient. The invention is useful for stimulating a T cell response in a
CC human patient, as probe or primer for nucleic acid hybridisation, to
CC selectively form duplex molecules with complementary stretches of the
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
CC length gene from a suitable library, and to direct expression of a
CC polypeptide in appropriate host cells. The composition is useful in
CC prophylactic or therapeutic applications and for the treatment of cancer,
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
CC associated malignancies. The invention is useful in gene therapy. The
CC present sequence is human cDNA for the clone HICD-CT_His_coding_region.
XX
SQ Sequence 1767 BP; 381 A; 521 C; 529 G; 336 T; 0 other;

Alignment Scores:
Pred. No.: 4, 72e-67 Length: 1767
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 24

SEQ4 (1-266) x AAD32744 (1-1767)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 946 CAGAAATGAGGACTGGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
DB 1006 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
DB 1066 TTCCTTCCTCCAGACCTTGGCCCCGGGGCGCTGGGGGCGATGGTCCACACAGCAGCCGAGC 1125
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB 1126 TCATCTACGAGGAGTGGGGTGGGACCTGACATAGGGCTGGAGCCCTCTGAGAGGAG 1185
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB 1186 GCCCCAGGCTCCACTGGCACCCTCCCAAGGGGCTGGCTCCGATGATTTGATGGTGAC 1245
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
DB 1246 CTGGGAATGGGGGAGCCCAAGGGGCTGCAAGGCTCCCAACACATGACCCAGCCCTCTA 1305
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140

DB 1306 CAGCGGTACAGTACAGGACCCACAGTACCCCTGCTGAGACTGATGGCTACGTTGCC 1365
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
DB 1366 CCCCTGACCTGACGCCCCCAGCCTGAATATGTGAACACGAGGAGTGTTCGGCCCCAGGCC 1425
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
DB 1426 CCTTCGCCCCGAGAGGGGCCCTCTGCTGCTGCCGACCTGCTGGTGGCCACTCTGGAAGG 1485
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
DB 1486 CCCAAGACTCTCTCCCCAGGGAAGAATGGGCTCGTCAAAGAGCTTTTTCCTTTGGGGGT 1545
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
DB 1546 GCCGTGGAGAACCCGAGTACTTGACACCCAGGAGGAGGAGTGGCCCTCAGCCCAACCT 1605
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTyrTyrTyrTyrTyrTyrTyr 240
DB 1606 CCTCTGCTTTCAGCCCCAGCCTTCGACAACCTCTATTACTGGGACGAGGCCACACAGAG 1665
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
DB 1666 CGGGGGGCTCCACCCAGCACCTTCAAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 1725
QY 261 GlyLeuAspValProVal 266
DB 1726 GGTCTGGAGTGGCAGTG 1743
RESULT 3
AAD32747
ID AAD32747 standard; cDNA; 1773 BP.
XX
AC AAD32747;
XX
DT 01-JUL-2002 (first entry)
XX Human cDNA for the clone HICD_in_pPDM_coding_sequence.
XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
XX human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..1770
XX /*tag= a
XX /product= "Human protein encoded by cDNA for the clone
XX HICD_in_pPDM_coding_sequence"
XX WO200214503-A2.
XX 21-FEB-2002.
XX 14-AUG-2001; 2001WO-US41733.
XX 14-AUG-2000; 2000US-225152P.
XX 28-SEP-2000; 2000US-236428P.
XX 21-FEB-2001; 2001US-270520P.
XX (CORI-) CORIXA CORP.
XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
PI McNeill PD, Vedvick TS;
XX WPI: 2002-280758/32.
DR P-PSDB; AAE20481.
XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,
XX prevention and diagnosis of cancer, preferably breast cancer
XX

Example 5; Page 119; 129pp; English.

PS The invention relates to an isolated Her-2/Neu polypeptide composition
XX effective for eliciting an immune response. The invention is useful for
CC eliciting an immune response in a patient, where the patient is human
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
CC The composition is useful for the therapy and diagnosis of cancer.
CC Preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
CC and other compositions for the diagnosis, prevention and treatment of
CC human malignancies, for stimulating and/or expanding T cells specific for
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
CC patient. The invention is useful for stimulating a T cell response in a
CC human patient, as probe or primer for nucleic acid hybridisation, to
CC selectively form duplex molecules with complementary stretches of the
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
CC length gene from a suitable library, and to direct expression of a
CC polypeptide in appropriate host cells. The composition is useful in
CC prophylactic or therapeutic applications and for the treatment of cancer,
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
CC associated malignancies. The invention is useful in gene therapy. The
CC present sequence is human cDNA for the clone HICD_in_ppdm_coding_
XX sequence.

SQ Sequence 1773 BP; 383 A; 528 C; 530 G; 332 T; 0 other;

Alignment Scores:

Pred. No.: 4.73e-67 Length: 1773
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ4 (1-266) x AAD32747 (1-1773)

QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 970 CAGAATGAGGACTGGGCCCGCAGCGACTGCTTGGACAGCACCTTACCGCTCACTGCTG 1029
QY 21 GluAspAspMetGlyAspLeuValAlaGluClyTyrLeuValProGlnGlnGly 40
DB 1030 GAGACGATGACATGGGGGACCTGGTGATGCTGAGAGATATCTGGTACCCACGAGGC 1089
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
DB 1090 TTCTTCTGTCAGACCTGCGCGCGCTGGGCGCATGCTCCACACAGCCAGCCGAGC 1149
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB 1150 TCATCTACACGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 1209
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB 1210 GCCCCAGGCTCTCACTGGCACCCCTCGAAGAGGGCTGGCTCCGATGATTTGATGGTGAC 1269
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
DB 1270 CTGGGAATGGGCGAGCCAGGCGCTGCAAGGCTCCCCACACATGACCCCGCCCTCTA 1329
QY 121 GluArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
DB 1330 CAGCGGTACAGTGGAGCCCGACAGTACCCCTGCTGAGACTGATGGCTACGTGTC 1389
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
DB 1390 CCCTGACCTGAGCCCGCCAGCTGATATGTGAACAGCCAGCATGTTCGGCCCGCCGCC 1449
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
DB 1450 CCTTCGCCCGGAGGGGCCCTCTGCTGCTGCCGACCTGCTGGTGGCCACTCTGGAAGG 1509
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200

Db 1510 CCCAAGACTCTCTCCCGAGGAGAAATGGGGTGGTCTCAAGACGCTTTTGGCTTTGGGGGT 1569
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 1570 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGAGCTGCCCTCAGCCCCCCT 1629
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 1630 CCTCTCTGCCCTCAGCCCGAGCTTCGACAACTCTATTACTGGGACGAGGACCCACAGAG 1689
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 1690 CGGGGGGCTCCACGACCTTCAAGGGGACCTACGCGAGAGACCCAGAGTACCTG 1749
QY 261 GlyLeuAspValProVal 266
Db 1750 GGTCTGGACGTGCCAGTG 1767
RESULT 4
AAT40739
ID AAT40739 standard; cDNA; 3768 BP.
XX
AC AAT40739;
XX
DT 01-JAN-1997 (first entry)
XX
DE HER-2/neu oncogene.
XX
KW HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
KW breast cancer; ovary cancer; colon cancer; lung cancer;
KW prostate cancer; genetic immunisation; tumour; vaccine; vector;
KW ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3765
FT FT /*tag= b
FT FT /note= "nucleotides 2026-3765 (claim 1) code for
FT FT HER-2/neu intracellular domain"
XX
PN W09630514-A1.
XX
PD 03-OCT-1996.
XX
PE 28-MAR-1996; 96WO-US01689.
XX
PR 31-MAR-1995; 95US-0414417.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Cheever MA, Disis ML;
XX
DR WPI; 1996-455361/45.
DR P-PSDB; AAW01111.
XX
PT DNA encoding HER-2-neu polypeptide(s) - used for prevention or
PT treatment of malignancies with which the HER-2/neu oncogene is
PT associated
XX
PS Claim 1; Page 49-56; 71pp; English.
XX
CC Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or
CC c-erbB2) protein (AAW01111). The oncogene is overexpressed in various
CC cancers, including breast, ovarian, colon, lung and prostate, and
CC appears to induce malignancies through quantitative mechanisms that
CC result from increased or deregulated expression of an essentially
CC normal gene product. Nucleotides 2026-3765 of the cDNA sequence
CC code for the intracellular domain (lys676-Val1255) of the HER-2/neu
CC protein, which is useful for immunisation against malignancy.
CC Nucleic acids can be used to direct expression of the intracellular
CC domain in transformed host cells, or are used, alone or in a viral
CC vector, for genetic immunisation of an animal.

XX SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 9,23e-67 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

SEQ4 (1-266) x AAX01912 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 2968 CAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
QY 21 GluAspAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
DB 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTATCCCGCAGCAGGC 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60
DB 3088 TTCCTTGTCCAGACCTTGGCCCGGCGCTGGGGCGATGGTCCACAGCAGCCAGCCGACG 3147
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB 3148 TCATCTACAGGAGTGGCGTGGGACCTGACACTAGGCTGGAGCCCTCTGAAGAGGAG 3207
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB 3208 GCGCCCGAGGCTCCACTGGCACCCCTCCGAAGGGGCTGCGATGTATTTGATGGTGAC 3267
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
DB 3268 CTGGGAATGGGGGACCCAGGGGCTGCARAGCCTCCACACATGACCCAGCCCTCTA 3327
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
DB 3328 CAGCGGTACAGTGGAGGACCCACACAGTACCCCTGCCCTCTGAGACTGATGGCTACGTC 3387
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
DB 3388 CCCCTGACCTGCAGCCCGCCAGCCTGAATATGTGAACCCAGCCAGATGTTCCGGCCG 3447
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
DB 3448 CCTTCGCCCGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGGCAGCTCTGGAAGG 3507
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
DB 3508 CCCAAGACTCTCTCCCGAGGGAAGTGGGGTGGTCAAGAGCTTTTTCCTTTGGGGGT 3567
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
DB 3568 GCGGTGGAGAACCCCGACTACTTACACCCCGAGGAGGAGTGCCTCCAGCCACACCT 3627
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTyrAspGlnAspProProGlu 240
DB 3628 CCTCTGCTTCAGCCCGACCTTCGACCAACCTCTATTACTGGACGAGCCAGCCAGCAG 3687
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
DB 3688 CGGGGGGCTCCACCCAGCACCTTCAAGGGGACACCTACGGCAGAGAACCCAGAGTAC 3747
QY 261 GlyLeuAspValProVal 266
DB 3748 GGTCTGGACGTGCCAGTG 3765

RESULT 5

AAX01912

ID AAX01912 standard; DNA; 3768 BP.

XX

AC AAX01912;
XX 21-APR-1999 (first entry)
XX Human HER-2/neu oncogene DNA.
XX HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
KW malignancy; treatment; tumour; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 1..3768
FT CDS
FT /*tag= a
FT /product= "HER-2/neu"
FT /note= "oncogene"
FT misc_feature 2026..3765
FT /*tag= b
FT /note= "region which elicits immune response"
XX US5869445-A.
XX 09-FEB-1999.
XX 01-APR-1996; 96US-0625101.
XX 01-APR-1996; 96US-0625101.
PR 17-MAR-1993; 93US-0033644.
PR 12-AUG-1993; 93US-0106112.
PR 31-MAR-1995; 95US-0414411.
XX (UNIW) UNIV WASHINGTON.
XX Cheever MA, Disis ML;
XX WPI; 1999-152835/13.
DR P-PSDB; AAW92406.
XX Use of HER-2/neu polypeptides - for eliciting an immune response to
PT an HER-2/neu associated malignancy, particularly for treating or
PT preventing tumours
XX Claim 1a; Column 23-32; 26pp; English.
XX This sequence encodes the human HER-2/neu oncogene protein. A fragment
CC of this protein is used in a method for eliciting or enhancing an immune
CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
CC B cells to produce an immune response to the HER-2/neu protein. The
CC method can be used for immunisation against a malignancy in which the
CC HER-2/neu oncogene is associated and in the treatment of an existing
CC tumour, or to prevent tumour occurrence or recurrence.
XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
Alignment Scores:
Pred. No.: 9,23e-67 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0
SEQ4 (1-266) x AAX01912 (1-3768)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 2968 CAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
QY 21 GluAspAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
DB 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTATCCCGCAGCAGGC 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60

```

Db      3088  TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGCTCCACACAGCCAGCCGAGC 3147
Qy      61  SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db      3148  TCATCTACCAAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3207
Qy      81  AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db      3208  GCCCCAGGCTCTCACTGGCACCTCCGAAGGGGCTGGCTCCGATGATTTGATGTGAC 3267
Qy      101  LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db      3268  CTGGGAATGGGGCAGCCAAAGGGGCTGCAAAAGCTCCCCACACATGACCCAGCCCTCTA 3327
Qy      121  GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db      3328  CAGCGGTACAGTGAGGACCCACAGTACCCCTGCCCTCTCAGACTGATGGCTAGTTGCC 3387
Qy      141  ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db      3388  CCCCTGACCTGCAGCCGCCAGCTGATATGTGAACAGCCAGATGTTGCGCCCCAGCCC 3447
Qy      161  ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db      3448  CCTTCGCCCGCAGAGGGCCCTCTGCCTGCTGCCCGACCTGCTGGTGCCACTCTGGAAGG 3507
Qy      181  ProIysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db      3508  CCCAAGACTCTCTCCCGCAGGAGAAATGGGGTCTGCTCAAGACGTTTTTGTCTTGGGGT 3567
Qy      201  AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db      3568  GCGGTGGAGAACCCGAGTACTTGACACCCCGAGGAGAGCTGCCCTCAGCCGCCCTT 3627
Qy      221  ProProAlaPheSerProAlaPheAspAsnLeuTyrTrpAspGlnAspProProGlu 240
Db      3628  CCTCCTGCCCTCAGCCCGAGCTTCGACACCTCTATTACTGGACCCAGGACCCACAGAG 3687
Qy      241  ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db      3688  CGGGGGGCTCCACCCAGCACCTTCAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3747
Qy      261  GlyLeuAspValProVal 266
Db      3748  GGTCTGCAGCTGCCAGTG 3765

RESULT 6
AAH23392
ID  AAH23392 standard; DNA; 3768 BP.
XX
AC  AAH23392;
XX
XX
AC  AAH23392;
XX
XX
DT  25-SEP-2001 (first entry)
XX
DE  Human HER-2/neu protein encoding DNA.
XX
KW  Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
XX  oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.
XX
OS  Homo sapiens.
XX
XX  Key Location/Qualifiers
XX  CDS 1..3768
XX  FT  /*tag= a
XX  FT  /product= "HER-2/neu protein"
XX
XX  WO200153463-A2.
XX
XX  26-JUL-2001.
XX
XX  19-JAN-2001; 2001WO-US01850.
XX

```

21-JAN-2000; 2000US-0177545.
(CORI-) CORIXA CORP.
Cheever MA, Hand-Zimmermann S;
P-PSDB; AAB85458.
WPI; 2001-476112/51.
DR
DR
XX
PT New antigen-presenting cells, useful as vaccines for eliciting or
enhancing an immune response to HER-2/neu protein, particularly useful
for treating or preventing cancer, e.g. breast cancer -
XX
PS Claim 1; Page 41-46; 49pp; English.
XX
CC The invention provides an isolated antigen-presenting cell, which
expresses at least an immunogenic portion of a polypeptide that produces
an immune response to HER-2/neu protein. The antigen-presenting cells are
useful as vaccines for eliciting or enhancing an immune response to
HER-2/neu protein, particularly in treating or preventing malignancies in
which the HER-2/neu oncogene is associated. Specifically, these are
useful for treating or preventing cancer, e.g. breast cancer, ovarian,
colon, lung or prostate cancers. The present sequence represents a DNA
encoding the human HER-2/neu protein (also known as p185 or c-erbB2).
XX
SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 9,23e-67 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

SEQ4 (1-266) x AAH23392 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGATGAGGACATGGGGCCAGCCAGTCCCTTGACAGCAGCCTTCTACCTCAGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGAGCATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGTGTACCCAGCAGGC 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTGTCCAGACCTTGCCTGGGGGCTGGGGGATGCTCCACACAGCAGCCGAGC 3147
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3148 TCATCTACCAAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3207
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCCCCAGGCTCTCACTGGCACCTCCGAAGGGGCTGGCTCCGATGATTTGATGTGAC 3267
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGAATGGGGCAGCCAAAGGGGCTGCAAAAGCTCCCCACACATGACCCAGCCCTCTA 3327
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGGTACAGTGAGGACCCACAGTACCCCTGCCCTCTCAGACTGATGGCTAGTTGCC 3387
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3388 CCCCTGACCTGCAGCCGCCAGCTGATATGTGAACAGCCAGATGTTGCGCCCCAGCCC 3447
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3448 CCTTCGCCCGCAGAGGGCCCTCTGCCTGCTGCCCGACCTGCTGGTGCCACTCTGGAAGG 3507

QY 181 ProLysThrLeuSerProGlyLysAsnGlyValLysAspValPheAlaPheGlyGly 200
 Db CCCAGACTCTCTCCAGGGAAGAAATGGGTCGTCAAAGACGTTTTTGGGGGT 3567
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
 Db GCGGGGAGAACCCCGAGTACTTGACACCCAGGAGGAGTGCCTTCAGCCCAACCT 3627
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
 Db CCTCTGCTTCAGCCAGCCCTTCGACAACTCTATTAATCTGGACAGGACCCACAGAG 3687
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
 Db CGGGGGGCTCCAGCCAGCACCTTCAAGGGGACACCTAGCGCAGAGAACCCAGAGTACCTG 3747
 QY 261 GlyLeuAspValProVal 266
 Db GGTCTGGACGTGCCAGTG 3765
 RESULT 7
 AAD32743
 ID AAD32743 standard; DNA; 3768 BP.
 AC AAD32743;
 XX
 XX 01-JUL-2002 (first entry)
 XX Human Her-2/neu protein DNA.
 XX
 XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ds.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 1..3768
 FT /*tag= a
 FT /product= "Human Her-2/neu protein"
 FT misc_feature 2026..3765
 FT /*tag= b
 FT /note= "Intracellular domain"
 FT
 PN WO200214503-A2.
 XX
 XX 21-FEB-2002.
 XX
 XX 14-AUG-2001; 2001WO-US41733.
 XX
 XX 14-AUG-2000; 2000US-225152P.
 PR 28-SEP-2000; 2000US-236428P.
 PR 21-FEB-2001; 2001US-270520P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
 PI McNeill PD, Vedvick TS;
 XX
 XX WPI; 2002-280758/32.
 DR P-PSDB: AAE20479.
 XX
 XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,
 PT prevention and diagnosis of cancer, preferably breast cancer -
 XX
 XX Claim 9; Page 109-114; 129pp; English.
 XX
 XX The invention relates to an isolated Her-2/Neu polypeptide composition
 CC effective for eliciting an immune response. The invention is useful for
 CC eliciting an immune response in a patient, where the patient is human
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
 CC The composition is useful for the therapy and diagnosis of cancer,
 CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
 CC and other compositions for the diagnosis, prevention and treatment of

human malignancies, for stimulating and/or expanding T cells specific for
 Her-2/Neu polypeptide and for inhibiting the development of cancer in a
 patient. The invention is useful for stimulating a T cell response in a
 human patient, as probe or primer for nucleic acid hybridisation, to
 selectively form duplex molecules with complementary stretches of the
 entire Her-2/Neu gene or gene fragments of interest, to isolate a full
 length gene from a suitable library, and to direct expression of a
 polypeptide in appropriate host cells. The composition is useful in
 prophylactic or therapeutic applications and for the treatment of cancer,
 preferably for the immunotherapy of breast cancer and other Her-2/Neu-
 associated malignancies. The invention is useful in gene therapy. The
 present sequence is human Her-2/neu protein DNA.
 XX
 SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
 Alignment Scores:
 Pred. No.: 9,23e-67 Length: 3768
 Score: 1450.00 Matches: 266
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 SEQ4 (1-266) x AAD32743 (1-3768)
 QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 Db CAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACACACCTTCTACCGTCTACTGCTG 3027
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
 Db GAGGACGATGATGGGGGACCTGCTGAGGAGTATCTGGTACCCAGCAGGAGG 3087
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
 Db TTCTTCTGTCAGACCTTGCCTGGGGGCTGGGGGATGGTCCACACAGCAGCCGACG 3147
 QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
 Db TCATCTACCAGAGTGGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3207
 QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
 Db GCGCCAGGCTCTCCACTGGCACCTCCGAAAGGGGCTGGCTCCGATGTATTTGATGTGAC 3267
 QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
 Db CTGGGAATGGGGGAGCCAGGGGCTGCAAGAGCTCCCAACACATGACCCAGCCCTCTA 3327
 QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
 Db CAGCGGTACAGTGAGGACCCACAGTACCTCCCTCTGACACTGATGGTACCTGCC 3387
 QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
 Db CCCCTGACCTGCAGCCCGCCAGCCCTGAATATCTGAACACAGCAGAGTGTTCGCCCCAGCCC 3447
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
 Db CCTTCCGCCCGAGAGGGCCCTCTCTGCTCCCGGACCTGTGTGGTGGCCACTCTGGAAAGG 3507
 QY 181 ProLysThrLeuSerProGlyLysAsnGlyValLysAspValPheAlaPheGlyGly 200
 Db CCCAAGACTCTCTCCCCAGGGAAGAAATGGGTCGTCAAAGACGTTTTTGGGGGT 3567
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
 Db GCGGTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGTGCCTTCAGCCCAACCT 3627
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
 Db CCTCTGCTTCAGCCAGCCAGTTCGACAACTCTATTAATCTGGACAGGACCCACAGAG 3687

Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3688 CGGGGGCTCCACCCAGCACCTTCAAAGGGACACCTACGGCAGAACCCAGAGTACCTG 3747

Qy 261 GlyLeuAspValProVal 266
Db 3748 GGTCTGGACGTGCCAGTG 3765

RESULT 8
ID ABA92250 standard; cDNA; 3768 BP.
XX ABA92250;

XX 17-JUN-2002 (first entry)
XX Human Her-2/neu cDNA.

XX Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
KW receptor; human; gene therapy; gene; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
FT CDS 1..3768
FT /*tag= a
FT /product= "Her-2/neu"

XX WO200212341-A2.

XX 14-FEB-2002.

XX 03-AUG-2001; 2001WO-US24283.

XX 03-AUG-2000; 2000US-0632507.

XX (CORI-) CORIXA CORP.
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Cheever MA, Gheysen D;

XX WPI; 2002-241743/29.

XX P-PSDB; AAMS1143.

XX Her-2/neu fusion protein for treating or preventing cancer by eliciting
XX or enhancing an immune response to the protein, has Her-2/neu
XX extracellular domain fused to Her-2/neu intracellular or
XX phosphorylation domain

XX Disclosure; Fig 15; 141pp; English.

XX The present sequence is that of human Her-2/neu oncogene cDNA.
XX The cDNA encodes Her-2/neu (p185), an oncogenic self protein and
XX target for anti-cancer vaccines. The Her-2/neu gene is amplified
XX and p185 is overexpressed in a variety of cancers, including breast,
XX ovarian, colon, lung and prostate cancer. Her-2/neu is a member
XX of the tyrosine kinase family of receptor-like glycoproteins.
XX Its overexpression correlates with a poor prognosis in breast and
XX ovarian cancers. The invention provides Her-2/neu fusion
XX proteins, nucleic acids encoding them, viral vectors, and vaccines
XX comprising the fusion proteins or nucleic acid molecules. In
XX preferred fusion proteins, the extracellular domain of a Her-2/neu
XX protein is fused to a Her-2/neu intracellular domain or
XX phosphorylation domain (or its DeltaPD fragment). An immune
XX response to Her-2/neu protein is elicited or enhanced by
XX administering the fusion protein in the form of a vaccine, or by
XX transfecting cells of an animal ex vivo with a nucleic acid
XX encoding the fusion protein, and delivering the transfected cells
XX to the animal. The fusion proteins, nucleic acids, and isolated
XX specific T-cells are useful for inhibiting the development of a
XX cancer, especially breast, ovarian, colon, lung or prostate cancer
XX in a patient. T cells that specifically react with a Her-2/neu
XX fusion protein can be used to remove tumour cells from a sample in

CC order to inhibit the development of cancer in a patient.
XX
SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 9,23e-67 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ4 (1-266) x ABA92250 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACAGCACCTTCTACGGCTCACGTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCACGAGGGC 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTCTCTCCAGACCTGCCCCGGCGCTGGGGGCATGGTCCACCAGAGCCAGCCGACG 3147
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3148 TCATCTACCAGGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGGCCCTCTGAAGAGGAG 3207
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCCCCCAGGTCTCCACTGGCACCTCCGAAGGGGCTGGCTCCCATGTATTTGATGGTGAC 3267
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGAATGGGGCAGCAAGGGGCTGCAAAAGGCTCCCCACACATGACCCAGCCCTCTA 3327
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGGTACAGTGAGGACCCACAGTACCCCTCCCTCTGAGACTGATGGCTACGTTGCC 3387
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3388 CCCTTGACCTGCAGCCCCCAGCCCTGAATATGTGAACCCAGCCAGATGTTGGCCCCCAGCCC 3447
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3448 CCTTCGCCCGGAGAGGGGCTCTGCTGCTGCCGACCTGCTGGTGCCTCTGGAAGG 3507
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3508 CCCAAGACTCTCTCCCCAGGGAAGATGGGGTGGTCAAGACGTTTGTGCTTTGGGGGT 3567
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 3568 GCCGTGGAGAACCCCGAGTACTGTACACCCCGAGGAGGAGCTGCCCTCTAGCCCCCACCCT 3627
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3628 CCTCCTGCCTTCAGCCAGCCTTCGACAACCTCTATTACTGGGACCCAGGACCCACCAGAG 3687
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3688 CGGGGGGCTCCACCCAGCACCTTCAAAAGGGACACCTACGGCAGAGAACCAGAGTACCTG 3747
Qy 261 GlyLeuAspValProVal 266
Db 3748 GGTCTGGACGTGCCAGTG 3765

RESULT 9
ABK10730
ID ABK10730 standard; DNA; 3768 BP.

XX ABK10730;
 AC 05-JUN-2002 (first entry)
 DT Human Her-2/neu DNA.
 DE Human; Her-2/neu; gene; ds; cytostatic; haematological malignancy; CML;
 XX acute myelogenous leukaemia; AML; chronic myelogenous leukaemia; CLL;
 KW chronic lymphocytic leukaemia; myeloma; non-Hodgkin's lymphoma; MDS;
 KW Hodgkin's lymphoma; T cell therapy.
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH 1..3768
 FT /*tag= a
 TT /product= "Human Her-2/neu"
 XX
 XX WO200213847-A2.
 XX 21-FEB-2002.
 XX 13-AUG-2001; 2001WO-US25408.
 XX 14-AUG-2000; 2000US-0638280.
 PR 28-SEP-2000; 2000US-0675904.
 XX (CORI-) CORIXA CORP.
 PA
 XX
 XX Galger A, Cheever MA, Hand-zimmermann S;
 PI
 XX WPI; 2002-280741/32.
 DR P-PSDB; AAU77114.
 XX
 XX Inhibiting haematological malignancy development by administering
 PT polypeptide comprising immunogenic portion of Her-2/neu, polynucleotide
 PT encoding the polypeptide, or antigen presenting cells expressing the
 PT polypeptide -
 XX
 XX Disclosure; Page 66-71; 74pp; English.
 XX
 XX The invention relates to a method for inhibiting development of
 CC haematological malignancy in a patient by administering a polypeptide
 CC comprising an immunogenic portion of Her-2/neu or a polynucleotide
 CC encoding the polypeptide. Antigen presenting cells that express the
 CC protein can also be administered. The sequences are used for inhibiting
 CC development of haematological malignancy such as acute myelogenous
 CC leukaemia (AML), chronic myelogenous leukaemia (CML), chronic lymphocytic
 CC leukaemia (CLL), MDS, myelomas, Hodgkin's lymphoma and non-Hodgkin's
 CC lymphoma. This sequence represents DNA encoding human Her-2/neu
 CC polypeptide.
 XX
 XX SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
 XX
 XX Alignment Scores:
 Pred. No.: 9,23e-67 Length: 3768
 Score: 1450.00 Matches: 266
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 SEQ4 (1-266) x ABK10730 (1-3768)
 QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 DB 2968 CAGATGAGGACTTTGGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
 DB 3028 GAGGACGATGATGATGGGGGACCTGTGTGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3087

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60
 DB 3088 TTTCTTCTCCAGACCTCCCGGGCGCTGGGGCGATGGTCCACACAGCAGCCGACG 3147
 QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGlyProSerGluGlu 80
 DB 3148 TCATCTACAGAGTGGGGTGGGACCTGACACACTAGGGCTGGAGCCCTCTCAAGAGGAG 3207
 QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyArg 100
 DB 3208 GCCCCAGGTCTCCACTGGCACCTCCGAAAGGGGCTGGCTCGATGATTTGATGGTGAC 3267
 QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
 DB 3268 GTGGGAATGGGGGAGCCCAAGGGGCTGCAAGCCCTCCCCACACATGACCCAGCCCTCTA 3327
 QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
 DB 3328 CAGCGGTACAGTGAGGACCCACAGTACCCCTGCCCCCTCTGAGACTGATGGCTACGTTGCC 3387
 QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
 DB 3388 CCCCTGACCTGCAGCCGCCAGCTCAATATGTGAACACAGCAGATGTTGGGGCCAGCC 3447
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
 DB 3448 CCTTCGCCCGAGAGGGGCCCTCTGCTGCTGCCGACCTGTGTGGTGCACCTCTGGAAGG 3507
 QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
 DB 3508 CCCAAGACTCTCTCCCCAGGGAAGTGGGGTCTCAAGAGCGTTTTTGGCTTTGGGGGT 3567
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
 DB 3568 GCCGTGGAGAACCCCGAGTACTTGACCCCCAGGAGGAGCTGCCCTCAGCCCCACCCCT 3627
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
 DB 3628 CCTCTGCTTTCAGCCAGCCTTCGACAACCTCTATTACTGGGACCAGGACCCACAGAG 3687
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
 DB 3688 CGGGGGGTCCACCCAGCACCTTCAAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3747
 QY 261 GlyLeuAspValProVal 266
 DB 3748 GGTCTGGACGTGCCAGTG 3765
 RESULT 10
 AAQ46083
 ID AAQ46083 standard; cDNA; 4299 BP.
 XX
 XX AC AAQ46083;
 XX
 XX DT 25-MAR-2003 (updated)
 DT 07-FEB-1994 (first entry)
 XX
 XX DE Sequence encoding a c-erbB-2 tumour antigen.
 XX
 XX KW Tumour antigen; c-erbB-2; glycoprotein; ss.
 XX
 XX OS Homo sapiens.
 XX
 XX FH Key Location/Qualifiers
 FT CDS 1..4299
 FT /*tag= a
 XX
 XX PN WO9316185-A2.
 XX
 XX PD 19-AUG-1993.
 XX
 XX PF 05-FEB-1993; 93WO-US01055.
 XX
 XX

PR 06-FEB-1992; 92US-0831967.
 XX (CETU) CETUS ONCOLOGY CORP.
 PA (CREA-) CREATIVE BIOMOLECULES INC.
 XX
 PI Houston LL, Huston JS, Oppermann H, Ring DB;
 XX WPI; 1993-272889/34.
 DR P-PSDB; AAR39568.
 XX
 PT New single chain Fv polypeptide binding to C-erbB-2 tumour
 PT antigen - for imaging or treating breast or ovarian cancer etc.
 XX
 PS Disclosure; pages 48-54; 87pp; English.
 XX
 CC c-erbB-2 refers to a protein antigen expressed on the surface of
 CC tumour cells, such as breast and ovarian tumour cells, which is an
 CC approx. 200,000 mol. wt. acidic glycoprotein having an isoelectric
 CC pt. of about 5.3 (see AAQ46083, AAR39568).
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 4299 BP; 882 A; 1307 C; 1266 G; 844 T; 0 other;

Alignment Scores:
 Pred. No.: 1.04e-66 Length: 4299
 Score: 1450.00 Matches: 266
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

SEQ4 (1-266) x AAQ46083 (1-4299)

QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 DB CAGAAAGAGGACTGGGCGCCAGCCAGTCCCTTGGACAGACCTTCTACCGCTCACTGCTG 3027
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
 DB GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGC 3087
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
 DB TTCTTCTGTGTCAGACCTGCGCGGGCGCTGGGGCATGCTCCACACAGCAGCCGCGAGC 3147
 QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
 DB TCATCTACCGAGGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3207
 QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
 DB GCGCCAGGTCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGTATTTGATGGTGAC 3267
 QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
 DB CTGGGAATGGGGGAGCCCAAGGGGCTCAAAAGGCTCCACACATGACCCAGCCCTCTA 3327
 QY 121 GluArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
 DB CAGCGGTACAGTGGAGGCCCAAGTACCTCCCTCTGAGACTGATGGCTACGTGGCC 3387
 QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
 DB CCGCTGACCTGGACCCCGCCAGCTGAATATGTACACGACCATGTTCCGCCCCAGCCC 3447
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
 DB CCTTCGCCCGGAGAGGGCCCTCTGCTGCTGCTGCCGACCTGCTGGTGCACCTCTGGAAGG 3507
 QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
 DB CCAAGACTCTCTCCCGAGGGAAGTGGGGTCTCAAAAGACGTTTTTGTGCTTTGGGGT 3567

QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
 DB GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGTGCCTCTCAGCCCAACCT 3627
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
 DB CCTCCTGCCTTCAGCCCGAGCTTCGACACCTCTATTACTGGGACAGGACCCACAGAG 3687
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
 DB CGGGGGGCTCCACCCAGCCTTCAAAAGGACACCTACGCGAGAGAACCCAGAGTACCTG 3747
 QY 261 GlyLeuAspValProVal 266
 DB GGTCTGGACGTGCCAGTG 3765

RESULT 11
 AAA14812
 ID AAA14812 standard; cDNA; 4472 BP.
 XX
 AC AAA14812;
 XX
 DT 08-AUG-2000 (first entry)
 XX
 DE cDNA encoding the SPLICE erbb-2 receptor protein.
 XX
 KW SPLICE erbb-2 receptor protein; cell transformation disorder; cancer;
 KW tumor cell proliferation; tissue degeneration; arthropathy;
 KW bone resorption; inflammatory disease; degenerative disorder;
 KW wound healing; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 175..3942
 FT /*tag= a
 FT /product= "SPLICE erbb-2 protein"
 FT
 PN WO200020579-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 01-OCT-1999; 99WO-CA00912.
 XX
 PR 02-OCT-1998; 98US-0165192.
 XX
 PA (UYMC-) UNIV MCMASTER.
 XX
 PI Muller WJ, Siegel PM;
 XX
 DR WPI: 2000-303768/26.
 DR P-PSDB; AAY84780.
 XX
 PT Nucleic acid encoding an erbb 2 receptor protein designated SPLICE
 PT erbb-2, inhibitors of the protein are useful for treatment of cancer -
 XX
 PS Claim 4; Fig 1; 60pp; English.
 XX
 CC The present sequence encodes a SPLICE erbb-2 receptor protein. The
 CC protein has an in-frame deletion of 16 amino acids, 2 of which are
 CC conserved cysteine residues, compared to the unspliced protein. The
 CC erbb-2 polynucleotide is used to construct probes for detecting
 CC disorders of cell transformation such as cancer. Antibodies to the
 CC protein may be used to detect SPLICE erbb-2 in a sample. Agents
 CC (e.g. antisense oligonucleotides) which inhibit the expression of
 CC SPLICE erbb-2 are useful for reducing tumor cell proliferation and
 CC treating cancer. Substances which stimulate SPLICE erbb-2 are useful
 CC for treating conditions involving damaged cells including conditions
 CC in which degeneration of tissue occurs, such as arthropathy, bone
 CC resorption, inflammatory diseases, degenerative disorders of the
 CC central nervous system and wound healing.
 XX
 SQ Sequence 4472 BP; 902 A; 1383 C; 1328 G; 859 T; 0 other;

Alignment Scores:																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	</
-------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

21-OCT-2002 (first entry)
Human tumour antigen ErbB-2 DNA.
Tumour antigen; human; vaccine; cellular immune response; immunogen;
cancer; tumour; ErbB-2; ds.
Homo sapiens.
OS
US6287569-B1.
11-SEP-2001.
06-APR-1998; 98US-0056105.
10-APR-1997; 97US-043467P.
(REGC) UNIV CALIFORNIA.
Kipps TJ, Wu Y;
WPI; 1998-583198/49.
Generating cellular immune response in patient to target protein -
comprises introducing vector with nucleotide sequence encoding
immunogen comprising protein processing signal into cell of patient
Disclosure; Column 81-86; 61pp; English.
This invention describes a novel method for generating a cellular immune
response in a patient to a target protein or its fragment. The method
involves introducing a vector containing a nucleotide sequence encoding
a chimeric immunogen comprising a protein processing signal and the
target protein or its fragment. The immunogen is produced by the cells
and processed so that the target protein or its fragment is presented to
the patients immune system and a cellular immune response is initiated.
The method and vectors can be used as a form of vaccination and could be
used to generate a cellular immune response in patients to, e.g.
cancerous tumours. The cellular immune response is the predominant immune
response in the patient. This sequence represents a DNA fragment which
encodes the human tumour antigen ErbB-2 described in the method of the
invention.
Note: The information in this spec has been previously disclosed in
WO199845444 however this spec contained no sequence information.
XX
SQ Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other;

Alignment Scores:
Pred. No.: 1.07e-66 Length: 4473
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

SEQ4 (1-266) x ABQ76220 (1-4473)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 3142 CAGAATGAGGACTTGGGCCACCCAGTCCCTTGGACAGCACCTTACCGCTCACTGCTG 3201
QY 21 GluAspAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
DB 3202 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
DB 3262 TTCCTCTGTCCAGACCTTCGCCCGGGCGCTGGGGCATGGTCCACACAGGCACCGCAGC 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB 3322 TCATCTACGAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381

Db 3802 CCTCCTGCTTCAGCCAGCCTTCGACAACTCTATTACTGGACAGGACCCACAGAG 3861
 Qy 241 ArgGlyAlaProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
 Db 3862 CGGGGGCTCCACCCAGCACCTTCAAGGGACACTACGGCAGAGAACCCAGAGTACCTG 3921
 Qy 261 GlyLeuAspValProVal 266
 Db 3922 GGTCTGGACGTGCCAGT 3939
 RESULT 14
 ABZ34969
 ID ABZ34969 standard; cDNA; 4473 BP.
 AC ABZ34969;
 XX
 DT 05-FEB-2003 (first entry)
 XX
 DE Human gene expression profile polynucleotide SEQ ID NO 81.
 XX
 KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
 KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
 KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
 KW gene expression; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200274979-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 20-MAR-2002; 2002WO-US08456.
 XX
 PR 20-MAR-2001; 2001US-276947P.
 XX
 PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
 PI Wan J, Wang Y;
 XX
 DR WPI: 2002-740862/80.
 XX
 PT New gene expression profile generated from primary, endothelial,
 PT epithelial, and muscle cell types, useful for identifying disease
 PT pathologies involving alterations of gene expression, e.g. cancer
 XX
 PS Claim 3; Page 274-276; 850pp; English.
 XX
 CC The invention relates to a gene expression profile comprising one or more
 CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
 CC is a coronary artery endothelium, umbilical artery or vein endothelium,
 CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
 CC endothelium, myometrium microvascular endothelium, keratinocyte
 CC epithelium, bronchial epithelium, mammary epithelium, prostate
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
 CC small airway epithelium, renal epithelium, umbilical artery smooth
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
 CC osteoblasts or prostate stromal cell. The gene expression profile is used
 CC for determining the level of RNA expression for a sample, determining the
 CC phenotype of a cell and distinguishing cell types. The gene or a protein
 CC expression profile is useful in identifying disease pathologies
 CC involving alterations of gene expression. The assessment of expression
 CC profiles may provide meaningful information with respect to tumour type
 CC and stage, treatment methods, and prognosis. The gene or protein
 CC expression profile may also be used for creating microarrays. The
 CC microarray is useful for genetic and physical mapping of genomes, DNA
 CC sequencing, genetic or medical diagnosis, genotyping of organisms,
 CC confirming cell or tissue identifications and in identifying promising
 CC antibiotics, antiviral or antifungal agents.
 XX

Seq Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,07e-66 Length: 4473
 Score: 1450.00 Matches: 266
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 SEQ4 (1-266) x ABZ34969 (1-4473)
 Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 Db 3142 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTACCGCTCACTGCTG 3201
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
 Db 3202 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3261
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60
 Db 3262 TTCCTTCGTCGACAGCCCTGCCCCGGGGCTGGGGCATGGTCCACACAGGACCCAGCCAGC 3321
 Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
 Db 3322 TCATCTACCAGAGTGGGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381
 Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
 Db 3382 GCCCCAGGCTCTCCACTGGCACCTCCGAAGGGCTGGCTCGATGATGATGATGATGATG 3441
 Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
 Db 3442 CTGGAAATGGGGGAGCCCAAGGGCTGCAAGGCTCCCAACATGACATGACCCAGCCCTCTA 3501
 Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
 Db 3502 CAGCGGTACAGTGAGGACCCACAGTACCCCTGCTGAGACTGATGGCTACGCTTGC 3561
 Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
 Db 3562 CCCCTGACCTGACGCCCTGAAATATGTGAACACAGCCAGATGTTGCGGCCCCAGCC 3621
 Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
 Db 3622 CCTTCGCCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGGCCACTCTGGAAGG 3681
 Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValLysAspValPheAlaPheGlyGly 200
 Db 3682 CCCAAGACTCTCTCCCGCAGGAGAAATGGGGTGGTCTCAAGAGAGTCTTTTTCCTTTGGGGGT 3741
 Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
 Db 3742 GCCCTGGAGAACCCGAGTACTTGACACCCCGAGGAGGAGGTGCCCCCTCAGCCCCACGCT 3801
 Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
 Db 3802 CCTCTGCTTTCAGCCCCAGCCCTTCGACAACTCTATTACTGGGACGACGACCCACAGAG 3861
 Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
 Db 3862 CGGGGGGCTCCACCCAGCACCTTCAAGGGACACTTCAAGGGACGACCCAGAGAGTACCTG 3921
 Qy 261 GlyLeuAspValProVal 266
 Db 3922 GGTCTGGACGTGCCAGT 3939
 RESULT 15
 AAD38904
 ID AAD38904 standard; DNA; 4473 BP.
 XX
 AC AAD38904;

```

XX 23-SEP-2002 (first entry)
DT Human Her-2 DNA.
DE
XX
DE
XX
KW Human; Her-2; epidermal growth factor receptor 2; infection; cancer;
KW hyperproliferative disorder; prophylaxis; inflammation; antisense;
KW tumour; gene therapy; phosphorothioate backbone; gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 175..3942
XX FT /*tag= a
XX FT /product= "Human Her-2 protein"
XX
XX WO200222636-A1.
XX
XX 21-MAR-2002.
XX
XX 12-SEP-2001; 2001WO-US28572.
XX
XX 15-SEP-2000; 2000US-0663834.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Bennett CF, Cowser LM;
XX PI
XX WPI: 2002-471192/50.
XX DR P-PSDB; AAE24067.
XX
XX Novel antisense oligonucleotide which modulates the expression of Human
XX Epidermal Growth Factor receptor, Her2, is useful for treating tumors
XX inflammation or to prevent infection in humans .
XX
XX Example 13; Page 94-101; 116pp; English.
XX
XX The invention relates to antisense compounds targetted to a nucleic
XX acid molecule encoding Her2 (human Epidermal Growth Factor receptor 2)
XX that specifically hybridises with and inhibits the expression of Her2.
XX Antisense compounds of the invention are used for treating diseases or
XX conditions associated with Her2 such as hyperproliferative disorders
XX e.g. lung, breast, gastric, oesophageal, colon, bladder, salivary,
XX neural or cardiac cancer. They are also useful prophylactically e.g.
XX to prevent or delay infection, inflammation and tumour formation. The
XX invention is also used in gene therapy. The present sequence is human
XX Her-2 DNA.
XX
XX SQ Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other:
XX
XX Alignment Scores:
XX Pred. No.: 1.07e-66 Length: 4473
XX Score: 1450.00 Matches: 266
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 24 Gaps: 0
XX
XX SEQ4 (1-266) x AAD38904 (1-4473)
XX
XX 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
XX
XX 3142 CAGAAATGAGACACTGGGCCCGAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
XX
XX 21 GluAspAspMetGlyAspLeuValAspAlaGluClyTyrLeuValProGlnGlnGly 40
XX
XX 3202 GAGACCATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCCGACGAGGC 3261
XX
XX 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
XX
XX 3262 TTCTTCGTCCAGACCTGCCCGGGGCGCTGGGGCATGGTCCACACAGGACCGCAGC 3321
XX
XX 61 SerSerThrArgSerGlyGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 80

```

Search completed: October 15, 2003, 23:50:06
Job time : 277.648 secs

```

Db 3322 TCATCTACAGGAGTGGCGGTGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyValaGlySerAspValPheAspGlyAsp 100
Db 3382 GCGCCCGAGGTCTCCACTGGCACCCCTCCGAAGGGGCTGGCTCCGATGATTTGATGGTGAC 3441
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3442 CTGGGAATGGGGCAGCCAAAGGGCTGCAAGGCTCCCCACACATGAGCCCGCCCTCTA 3501
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3502 CAGCGGTACAGTGGAGGACCCACAGTACCCCTGCGCTCTGAGACTGATGGCTAGCTTGGC 3561
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3562 CCCCTGACCTGCAGCCCCAGCCCTGAATATGTGAACAGGACAGATGTTCGGCCCCAGCCC 3621
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3622 CCTTCGCCCCGAGAGGGGCCCTCTGCGCTGCTGCCGACCTGCTGGTGCCTCTGGAAGG 3681
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3682 CCCAAGACTCTCTCCCGAGGGAAGATGGGTGCTCAAGACGTTTTTGGCTTTGGGGGT 3741
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 3742 GCGGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCTTCAGCCCCACCT 3801
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3802 CCTCTGCGCTTCAGCCCGAGCCTTCGACAACTCTATTACTGGGACAGGACCCACAGAG 3861
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3862 CGGGGGGCTCCACCCAGCACCTTCAAGAGGACACCTACGGCAGAGAACCCAGAGTACCTG 3921
Qy 261 GlysLeuAspValProVal 266
Db 3922 GGCTGACCTGCCAGTG 3939

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 3830.63 Seconds
(without alignments)
2840.777 Million cell updates/sec

Title: SEQ4
Perfect score: 1450
Sequence: 1 QNEDLGASPLDSTFYSILL.....TFKGTPTAENPEYLGILDVVP 266

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2889711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model -DEV=xlp
-Q/cgn2_1/USPTO.spool_p/HOLLERAN480/runat_15102003_131912_20536/app_query.fasta_1.4685
-DB=GenEmbl -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 @CGN_1.1.22724 @runat_15102003_131912_20536 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -FGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	1450	100.0	1755	6	AX384609 Sequence
2	1450	100.0	1767	6	AX384607 Sequence
3	1450	100.0	1773	6	AX384610 Sequence
4	1450	100.0	3768	6	AR034479 Sequence
5	1450	100.0	3768	6	AX201817 Sequence
6	1450	100.0	3768	6	AX380923 Sequence
7	1450	100.0	3768	6	AX384604 Sequence
8	1450	100.0	3768	6	AX465456 Sequence
9	1450	100.0	4473	6	AR080259 Sequence
10	1450	100.0	4473	6	AR167390 Sequence
11	1450	100.0	4473	9	HSRBR2R
12	1442	99.4	1806	6	AX384608 Sequence
13	1442	99.4	3768	6	AX505114 Sequence
14	1442	99.4	3768	6	AX060704 Sequence
15	1442	99.4	3768	6	AX467229 Sequence
16	1442	99.4	3768	6	AX481438 Sequence
17	1442	99.4	4530	6	AR202597 Sequence
18	1442	99.4	4530	6	AR283481 Sequence
19	1442	99.4	4530	6	AX282577 Sequence
20	1442	99.4	4530	6	AX587649 Sequence
21	1442	99.4	4530	6	AX644071 Sequence
22	1442	99.4	4530	6	BD005474 Cellular
23	1442	99.4	4530	6	I21124 Sequence 9
24	1442	99.4	4530	6	I59745 Sequence 9
25	1442	99.4	4530	9	HUMHER2A
26	1442	99.4	9274	6	AX060703 Sequence
27	1358	93.7	161815	9	AC079199 Homo sapi
28	1358	93.7	168585	9	AC040933 Homo sapi
29	1358	93.7	198008	2	AC142197 Homo sapi
30	1350	93.1	13450	9	AB096612 Homo sapi
31	1350	93.1	16572	9	AB096614 Homo sapi
32	1350	93.1	20271	9	AB096613 Homo sapi
33	1350	93.1	30837	9	AY208911 Homo sapi
34	1291.5	89.1	3780	4	AB008451 Canis fam
35	1225.5	84.5	4062	10	HAMNEU
36	1208	83.3	3955	6	AX380924 Sequence
37	1208	83.3	3955	6	I21129 Sequence 14
38	1208	83.3	3955	6	I59750 Sequence 14
39	1208	83.3	3955	10	RNNEUR
40	1208	83.3	4727	10	AY116182 Rattus no
41	1205	83.1	2083	10	BC027080 Mus muscu
42	1205	83.1	2763	6	AX380942 Sequence
43	1205	83.1	2781	6	AX380944 Sequence
44	1205	83.1	3771	6	AX189662 Sequence
45	1205	83.1	3771	6	AX380925 Sequence

ALIGNMENTS

RESULT 1

```

AX384609
LOCUS AX384609 1755 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 6 from Patent WO0214503.
ACCESSION AX384609
VERSION AX384609.1 GI:19577810
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Hand-Zimmermann,S., Cheever,M.A., Foy,T.M., Lodes,M.J., Kalos,M.D.,
McNeill,P.D. and Vedvick,T.S.
TITLE Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 6 21-FEB-2002;
KEYWORDS CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..1755
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 376 a 517 c 531 g 331 t
ORIGIN
Alignment Scores:
Pred. No.: 7.4e-54 Length: 1755
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ4 (1-266) x AX384609 (1-1755)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 946 CAGAATGAGGACTTGGGGCCACGCCAGTCCCTTGGACAGACCTTCTACCGCTCACTGCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
DB 1006 GAGCAGCATGACATGGGGACCTGGTGATGCTCAGGAGTATCTGTGATCCACACGAGGC 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
DB 1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGTGTCACACAGCAGCCGCGAGC 1125
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB 1126 TCATCTACCAAGAGTGGGGACCTGACACTAGGGCTGGAGCCCTCTGGAAGAGGAG 1185
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB 1186 GCGCCAGGCTCCACTGGCACCTCCGAAAGGGCTGGCTCCGATGATTTGATGGTGAC 1245
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
DB 1246 CTGGGAATGGGGGAGCAAGGGCTGCAAGGCTCCACACATGATGATGATGATGATGATGAT 1305
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
DB 1306 CAGCGGTACAGTAGGAGACCCACAGTACCCCTGCGCTCTGAGACTGATGGCTGATGGTGC 1365
QY 141 ProLeuThrCysSerProGlnProGlnTyrValAsnGlnProAspValArgProGlnPro 160
DB 1366 CCCCTGACCTGCACCCCGAGCTGATATGTGACACAGCAGATGTTGGCCGCCACGCC 1425
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
DB 1426 CCTTCGCCCCGAGAGGGCCCTCTGCCTGCTGCCGACCTGTGTGTGCTGCTGCTGCTGCTG 1485
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
DB 1811 ProlLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200

```

```

Db 1486 CCCAAGACTCTCTCCCCAGGGAAGATGGGTCTCAAAGACGTTTTTGGCTTTGGGGGT 1545
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
DB 1546 GCCGTGGAGAACCCCGAGTACTTGACACCCAGGAGAGTGTCCCTTCACCCACCCT 1605
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
DB 1606 CCTCCTGCCITCAGCCCGAGCTTCGACAACTCTATTACTGGGACAGGACCCACAGAG 1665
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
DB 1666 CGGGGGCTCCACCGACGACCTTCAAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 1725
QY 261 GlyLeuAspValProVal 266
DB 1726 GGTCTGGACGTGCCAGTG 1743
RESULT 2
AX384607
LOCUS AX384607 1767 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 4 from Patent WO0214503.
ACCESSION AX384607
VERSION AX384607.1 GI:19577808
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Hand-Zimmermann,S., Cheever,M.A., Foy,T.M., Lodes,M.J., Kalos,M.D.,
McNeill,P.D. and Vedvick,T.S.
TITLE Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 4 21-FEB-2002;
FEATURES
source
Location/Qualifiers
1..1767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 381 a 521 c 529 g 336 t
ORIGIN
Alignment Scores:
Pred. No.: 7.45e-54 Length: 1767
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ4 (1-266) x AX384607 (1-1767)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 946 CAGAATGAGGACTTGGGGCCACGCCAGTCCCTTGGACAGACCTTCTACCGCTCACTGCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
DB 1006 GAGCAGCATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTGATCCACACGAGGC 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
DB 1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGTGTCACACAGCAGCCGCGAGC 1125
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB 1126 TCATCTACCAAGAGTGGGGACCTGACACTAGGGCTGGAGCCCTCTGGAAGAGGAG 1185
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB 1186 GCGCCAGGCTCCACTGGCACCTCCGAAAGGGCTGGCTCCGATGATTTGATGGTGAC 1245

```

QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
 Db 1246 CTGGGAATGGGGGAGCCAGGCGTCAAGAGCTCCACACATGATGAGCCCTCTA 1305
 QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
 Db 1306 CAGCGGTACAGTGGAGGACCCACACAGTACCCCTGCTGAGACTGATGGCTAGCTTGC 1365
 QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
 Db 1366 CCCCTGACTGACGCCCCCAGCCCTGATATGTGAACAGCCAGATGTTGGCCCGCCAGCCC 1425
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
 Db 1426 CCTTCGCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGCTGCTGCTGCTGCTGCTG 1485
 QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
 Db 1486 CCCAAGACTCTCTCCCGAGGGAAGTGGGTCGTCAAGAGCGTTTTCCTTTGGGGGT 1545
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
 Db 1546 GCCGTGGAGAACCCGAGTACTTGACACCCCGAGGAGGAGTGCCTCTACGCCACCCCT 1605
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
 Db 1606 CCTCTGCTTTCAGCCAGCTTTCGACAACTCTATTACTGGGACGAGGCCACCCAGAG 1665
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
 Db 1666 CGGGGGGTCCACCCAGCACCTTCAAGGGGACCTAGGGGAGAGAACCCAGAGTACCTG 1725
 QY 261 GlyLeuAspValProVal 266
 Db 1726 GGTCTGGAGTGGCCAGTG 1743
 RESULT 3
 LOCUS AX384610 1773 bp DNA linear PAT 19-MAR-2002
 DEFINITION Sequence 7 from Patent WO0214503.
 ACCESSION AX384610
 VERSION AX384610.1 GI:19577811
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Hand-Zimmermann,S., Cheever,M.A., Foy,T.M., Lodes,M.J., Kalos,M.D.,
 McNeill,P.D. and Vedvick,T.S.
 TITLE Compositions and methods for the therapy and diagnosis of
 her-2/neu-associated malignancies
 JOURNAL Patent: WO 0214503-A 7 21-FEB-2002;
 CORIXA CORPORATION (US)
 FEATURES
 Location/Qualifiers
 1..1773
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 383 a 528 c 530 g 332 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,47e-54 Length: 1773
 Score: 1450.00 Matches: 266
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 SEQ4 (1-266) x AX384610 (1-1773)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 Db 970 CAGAATGAGGACTTGGGCCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTACTGCTG 1029
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
 Db 1030 GAGACGATGACATGGGGGACCTGTGGTATCTGAGGAGTATCTGGTACCCAGCAGGCG 1089
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60
 Db 1090 TTCTTCTGTCCAGACCTTCCCGCGGCGCTGGGGCATGTCTCCACACAGCAGCCGACG 1149
 QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
 Db 1150 TCATCTACAGAGTGGCGTGGGACCTGACACTAGGCTGGAGCCCTCTGAAGAGGAG 1209
 QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAsnGlyAsp 100
 Db 1210 GCCCCCAAGTCTCCACCTGGCACCTTCCGAAAGGGCTGGCTCCGATGATTTGATGGTGAC 1269
 QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
 Db 1270 CTGGGAATGGGGGAGCCAGGCGTCAAGAGCTCCACACATGACCCCGACCCCTCTA 1329
 QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
 Db 1330 CAGCGGTACAGTGGAGGACCCACAGTACCCCTGCTGAGACTGATGGTACGTTGCC 1389
 QY 141 ProluthrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
 Db 1390 CCCCTGACCTGACGCCCCCAGCCCTGAATATGTGAACAGCAGATGTTTCGGCCCCAGCCC 1449
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
 Db 1450 CCTTCGCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGCACCTCTGGAAAGG 1509
 QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
 Db 1510 CCCAAGACTCTCTCCCGAGGGAAGTGGGTCGTCAAGAGACGTTTTTTCCTTTGGGGGT 1569
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
 Db 1570 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGTGGCCCTTCAGCCCCACCT 1629
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
 Db 1630 CCTCTGCTTACGCCCGCCAGCTTGCACACCTCTATTACTGGGACGAGCAGCCAGAG 1689
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
 Db 1690 CGGGGGGTCCACCCAGCACCTTCAAGGGGACACCTAGGGCAGAGAACCCAGAGTACCTG 1749
 QY 261 GlyLeuAspValProVal 266
 Db 1750 GGTCTGGACGTGCCAGTG 1767
 RESULT 4
 LOCUS AR034479 3768 bp DNA linear PAT 29-SEP-1999
 DEFINITION Sequence 1 from patent US 5869445.
 ACCESSION AR034479
 VERSION AR034479.1 GI:5950084
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 3768)
 AUTHORS Cheever,M.A. and Disis,M.L.
 TITLE Methods for eliciting or enhancing reactivity to HER-2/neu protein
 JOURNAL Patent: US 5869445-A 1 09-FEB-1999;
 FEATURES Location/Qualifiers
 1..3768
 /organism="unknown"

BASE COUNT	759 a	1171 c	1119 g	719 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	1.42e-53	Length:	3768		
Score:	1450.00	Matches:	266		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
SEQ4 (1-266) x AR034479 (1-3768)					
Qy	1	GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu	20		
Db	2968	CAGAATGAGGACTTGGCCCGGCGCTGGAGCAGACCTTCTACCGCTCAGTCTG	3027		
Qy	21	GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGly	40		
Db	3028	GAGGACGATGACATGGGGACCTGGTGATCTGAGGATATCTGGTACCCAGAGGC	3087		
Qy	41	PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer	60		
Db	3088	TTCTTCTGTCACAGACCTGCCCGGCGCTGGGGCATGTCCACACAGCCAGCCGAGC	3147		
Qy	61	SerSerThrArgSerGlyGlyAspLeuGluProSerGluGluGlu	80		
Db	3148	TCATCTACGAGGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGG	3207		
Qy	81	AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly	100		
Db	3208	GCCCCAGGCTCCACATGGCACCCTCGAAGGGCTGGCTCCGATGATTTGATGTGAC	3267		
Qy	101	LeuGlyMetGlyAlaAlaLysGlyLeuLysSerLeuProThrHisAspProSerProLeu	120		
Db	3268	CTGGGAATGGGCGAGCAGAGGGCTGCAAGGCTCCCAACATGACATGCCACCTCTA	3327		
Qy	121	GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValala	140		
Db	3328	CAGCGGTACAGTGGAGCCACACAGTACCCCTGCTGAGACTGATGCTACGTTGCC	3387		
Qy	141	ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro	160		
Db	3388	CCCTTGACCTGACGAGCCCGGAGCTGGAATATGTAACAGCAGCATGTTGCGCCAGCC	3447		
Qy	161	ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg	180		
Db	3448	CCTTCGCGCGAGAGGCCCTCTGCTGCTGCGGACCTGCTGGTGCACCTCTGGAAAG	3507		
Qy	181	ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly	200		
Db	3508	CCCAAGACTCTCTCCCGAGGAAGATGGGCTGCTCAAGACGCTTTTGGCTTTGGGGT	3567		
Qy	201	AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro	220		
Db	3568	GCGGTGAGAACCCCGAGTACTTGACACCCGAGGAGGAGCTGCCCTTCAGCCGCCACCT	3627		
Qy	221	ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu	240		
Db	3628	CTCTCTGCTTACGCCGCTTGCACACCTCTATTACTGGGACGAGCCACCCAGAG	3687		
Qy	241	ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu	260		
Db	3688	CGGGGGCTCCACCCAGCCTTCAAGGGACACCTAGGCGAGAGAACCCAGAGTACCTG	3747		
Qy	261	GlyLeuAspValProVal	266		
Db	3748	GGTCTGGACGTGCCAGTG	3765		
RESULT 5					
AX201817					
LOCUS	AX201817	3768 bp	DNA	linear	PAT 30-AUG-2001
DEFINITION	Sequence 1 from Patent WO0153463.				

ACCESSION	AX201817
VERSION	AX201817.1
KEYWORDS	GI:15391666
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1
AUTHORS	Cheever, M.A. and Hand-Zimmermann, S.
TITLE	Compounds and methods for prevention and treatment of her-2/ neu associated malignancies
JOURNAL	Patent: WO 0153463-A 1 26-JUL-2001;
FEATURES	CORIXA CORPORATION (US)
source	Location/Qualifiers
	1..3768
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
CDS	1..3768
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAC60126.1"
	/db_xref="GI:15391667"
	/translation="MELAALCRWGLLLALLPFGAASQTQVCTGTMKLRPLPASPETHLD MLRLHQGGVQVQGNLELYTLPTNASLSFLQIQEVQGYVIIAHNQVRVQPLQRLRV RGQLQEDNYALAVLNDGDLNNTPTVTPGASGGLRELQLSLTELKGGVLIQRLVQ LCYQDITLMDKIDPHKNNQALATLIDTNRSRACHPCSPMCKGSRGWSHSDCCSLRT VCAAGCARGKPLPTDCCHEQCAAGCTGPKHSDCLACILPHNSGICELCHPALVTVNT DTFESHPNPEGRTYFGASCTVACPYNYLSTDVGSCTLVCPHMQEVTAEQGTQCEKC SKPCARVCGLGMELHREVRVTSANIQEPAGCKIFGSLAFLPESFDGDPASNTAPL OPQLQVFTLEITGYLXISAWPDSLDPLSVFQNLQVIRLILHNGASLTLQGLGI SWGLRLSLRELGLALIHNTHLCHFVTPVMDQLFRNPHQALLHTANRPEDCEVGEG LACHOLCARGHCHGPGPTQVCNCSQFLRGQECVECRVLQGLPREYVNAHRCLPCHPE COPQNSVTCGCEPADQCVACAHVKDPPCVAPCPGSKVPLDSYMPLWCPDEGACQ PCPINTHSCVDLDDKGCAPAEQASPLTSIISAVGILLVVLGVVFGILLIKRQOKI RYKTMRLRLQETELVPLTPSGAMPNOAMRIKTELKRVKVLGSGAFVTVKGIWI PDGENKIPVAIKVRENTSPKANKEIDAEYVMAGVSPVSRLLGICUTSTVQLVT QLMYPGCLLDHVRNRLGSLDNLNMCQIAKMSYLEDVRLVLRHSDVMSYGVYVME NHVKTDFGLARLLDIDETEHADGGKVPKIMWALISILRRVTHQSDVMSYGVYVME LMTFGAKPDVGIPARETDLLEKGERLPOPICTIDVYIMVWKCMWIDSCRFREL VSEFSRMAPDQREVIQNEDELGPASPLDSTFYRSILLEDDMDGLVDAEYGLVPQGF FCPDPAAGAGVHRRHSSTRSGGDLTLGLEPSEEAAPRPLAPSEAGSGSDFDG DLGMAAGALQSLPTHDPSLPQYSEDTPVLPSEDTGVYAPLTCSPQPYVNPQDVR POPSPREGPLPAARPAAGATLERPKTISPKNGVVKDVFAGGAVENPEYLTPOGGA POPHPPPATSPAFDNLVYWDQDPPERGAPSTFKGTAEINPEYTLGLDVPV"
BASE COUNT	759 a 1171 c 1119 g 719 t
ORIGIN	
Alignment Scores:	
Pred. No.:	1.42e-53
Score:	1450.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
SEQ4 (1-266) x AX201817 (1-3768)	
Qy	1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db	2968 CAGAATGAGGACTTGGCCCGGCGCTGGAGCAGACCTTCTACCGCTCAGTCTG 3027
Qy	21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGly 40
Db	3028 GAGGACGATGACATGGGGACCTGGTGATCTGAGGATATCTGGTACCCAGAGGC 3087
Qy	41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db	3088 TTCTTCTGTCACAGACCTGCCCGGCGCTGGGGCATGTCCACACAGCCAGCCAGC 3147
Qy	61 SerSerThrArgSerGlyGlyAspLeuGluProSerGluGluGlu 80
Db	3148 TCATCTACGAGGATGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGG 3207

[illegible]

Db	3208	GC	CCCCCAGG	GCTCCAC	TGGCACC	CTCCGA	AGGGCTGG	CTCCGATG	TATTTGAT	GGTGAC	3267		
Qy	101	Leu	GlyMet	GlyAla	AlaLys	GlyLeu	GlnSer	LeuPro	ThrHis	AspPro	SerPro	Leu	120
Db	3268	CT	GGGAAT	GGGGCAG	CCCAAG	GGGCTG	CAAAAG	CCCTCC	CCACAC	ATGAC	CCACCC	AGCCCTCT	3327
Qy	121	Gln	ArgTyr	SerGlu	AspPro	ThrVal	ProLeu	LeuPro	SerGlu	ThrAsp	GlyTyr	ValAla	140
Db	3328	CAG	CGGTAC	AGTGAG	GACCCAC	AGTACC	CTGCCC	TGCGCT	CTGAG	ACTGAT	GCTACG	TTGGCC	3387
Qy	141	Pro	LeuThr	CysSer	ProGln	ProGlu	TyrVal	AsnGln	ProAsp	ValArg	ProGln	Pro	160
Db	3388	CC	CTGAC	CTGCAC	CCCCCAG	CCCTG	AATATG	TGA	ACCAG	CCAGATG	TTCTCG	GGCCCCAC	3447
Qy	161	Pro	SerPro	ArgGlu	GlyPro	LeuPro	ProAla	AlaArg	ProAla	GlyAla	ThrLeu	GluArg	180
Db	3448	CC	TGCGCC	CGCAG	AGGGCC	CTCTG	CCCTG	CTGCC	AGCTCT	GCGCAT	CTCTG	GGAAAG	3507
Qy	181	Pro	LysThr	LeuSer	ProGly	LysAsn	GlyVal	ValLys	AspVal	PheAla	PheGly	Gly	200
Db	3508	CCC	AAAGACT	CTCTCCC	CAGGCA	AGATGG	GTCTCA	AAACG	CTTTT	TGCTTT	TGCGGG	GT	3567
Qy	201	Ala	ValGlu	AsnPro	GluTyr	LeuThr	ProGln	GlyGly	AlaAla	ProGln	ProHis	Pro	220
Db	3568	GCG	CTGGAG	AACCC	CCAGTACT	TGAC	ACC	CCGAG	GAGCTG	CCCTC	TAG	CCCCAC	3627
Qy	221	Pro	ProAla	PheSer	ProAla	PheAsp	AsnLeu	TyrTrp	AspGln	AspPro	ProGlu	240	
Db	3628	CT	CTCGCC	TTACGCC	ACCGCT	TCGAC	CACTCT	TAT	TAC	TGGG	ACCA	CCAGAC	3687
Qy	241	Arg	GlyAla	ProPro	SerThr	PheLys	GlyThr	ProThr	AlaGlu	AsnPro	GluTyr	Leu	260
Db	3688	CGG	GGGGCT	CCACCC	AGCAC	CTTCAA	AGGAC	ACCTAC	GSCAG	AGAA	CCAC	CAGAGT	3747
Qy	261	Gly	LeuAsp	ValPro	Val	266							
Db	3748	GST	CTGGAC	GTGCC	AGTG	3765							
LOCUS	AR080259					4473 bp	DNA	linear					
DEFINITION	Sequence 1 from patent US 5968748.												
ACCESSION	AR080259												
VERSION	AR080259.1					GI:10006994							
KEYWORDS	Unknown.												
SOURCE	Unknown.												
ORGANISM	Unclassified.												
REFERENCE	1 (bases 1 to 4473)												
AUTHORS	Bennett,C.Frank., Lipton,A. and Witters,L.M.												
TITLE	Antisense oligonucleotide modulation of human HER-2 expression												
JOURNAL	Patent: US 5968748-A 1 19-OCT-1999;												
FEATURES	Location/Qualifiers												
source	1..4473												
BASE COUNT	902 a 1383 c 1329 g					859 t							
ORIGIN													
Alignment Scores:													
Pred. No.:	1.64e-53					Length:							
Score:	1450.00					Matches:							
Percent Similarity:	100.00%					Conservative:							
Best Local Similarity:	100.00%					Mismatches:							
Query Match:	100.00%					Indels:							
DB:	6					Gaps:							
SEQ4 (1-266) x	AR080259 (1-4473)												
Qy	1	Gln	AsnGlu	AspLeu	GlyPro	AlaSer	ProLeu	AspSer	ThrPhe	TyrArg	SerLeu	Leu	20
Db	3142	CAG	AATGAG	ACTTTG	GGCCAG	CCAGCAG	TCCCTT	GGACAG	CACCTT	CTAC	CGCTC	ACTGCTG	3201

```

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40
Db 3202 GAGGACCATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGC 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60
Db 3262 TTCTTCTGTCAGACCTGCCCGGGGCTGGGGCATGGTCCACACAGCCAGCCGAGC 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 80
Db 3322 TCATCTACAGGAGTGGCGTGGGACCTGACACTAGGCGTGGAGCCCTCTGAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 100
Db 3382 GCGGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3441
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 120
Db 3442 CTGGGATGGGCGAGCAAGGGCTGCAAGGCTCCCAAGCTCCACACATGACCCAGCCCTCTA 3501
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 140
Db 3502 CAGCGGTACAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3561
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 160
Db 3562 CCGCTGACCTGCAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3621
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 180
Db 3622 CCTTCGCGCCGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3681
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 200
Db 3682 CCAAGACTCTCTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3741
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3742 GCGGTGGAGAACCCGAGTACTTGACACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3801
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3802 CTCTCTGCTTACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3861
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr 260
Db 3862 CGGGGGGCTCCACCCAGCAGCTTCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3921
QY 261 GlyLeuAspValProVal 266
Db 3922 GGCTGGACGTGCAGTG 3939

```

```

RESULT 10
AR167390 4473 bp DNA linear PAT 17-DEC-2001
LOCUS
DEFINITION Sequence 26 from patent US 6287569.
ACCESSION AR167390
VERSION AR167390.1 GI:17903168
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 4473)
AUTHORS Kippes,T.J. and Wu,Y.
TITLE Vaccines with enhanced intracellular processing
JOURNAL Patent: US 6287569-A 26 11-SEP-2001;
FEATURES
source Location/Qualifiers
1..4473
/organism="unknown"
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:

```

```

Pred. No.: 1.64e-53 Length: 4473
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ4 (1-266) x AR167390 (1-4473)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGAATCAGGACTTGGGCGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40
Db 3202 GAGGACCATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGC 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60
Db 3262 TTCTTCTGTCAGACCTGCCCGGGGCTGGGGCATGGTCCACACAGCCAGCCGAGC 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 80
Db 3322 TCATCTACAGGAGTGGCGTGGGACCTGACACTAGGCGTGGAGCCCTCTGAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3382 GCGGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3441
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3442 CTGGGATGGGCGAGCAAGGGCTGCAAGGCTCCCAAGCTCCACACATGACCCAGCCCTCTA 3501
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3502 CAGCGGTACAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3561
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3562 CCGCTGACCTGCAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3621
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3622 CTCTGCGCGGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3681
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 200
Db 3682 CCAAGACTCTCTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3741
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3742 GCGGTGGAGAACCCGAGTACTTGACACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3801
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3802 CTCTCTGCTTACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3861
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3862 CGGGGGGCTCCACCCAGCAGCTTCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3921
QY 261 GlyLeuAspValProVal 266
Db 3922 GGCTGGACGTGCAGTG 3939

```

```

RESULT 11
HSEB2R
LOCUS
DEFINITION Human c-erb-B-2 mRNA.
ACCESSION X03363
VERSION X03363.1 GI:31197
KEYWORDS
cell surface glycoprotein; cellular oncogene; erb-B2 cellular
glycoprotein; growth factor receptor; neu cellular

```

oncogene; transmembrane protein; tyrosine kinase.
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 4473)
 Yamamoto,T., Ikawa,S., Akiyama,T., Samba,K., Nomura,N., Miyajima,N., Saito,T. and Toyoshima,K.
 Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor
 Nature 319 (6050), 230-234 (1986)
 86118663
 3003577
 Papewalis,J., Nikitin,A.Yu. and Rajewsky,M.F.
 G to A polymorphism at amino acid codon 655 of the human erbB-2/HER2 gene
 Nucleic Acids Res. 19 (19), 5452 (1991)
 92020265
 1681519
 The c-erb-B-2 protein shows similarity to the epidermal growth factor receptor.
 Location/Qualifiers
 1..4473
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /strain="cell line MKN-7"
 /db_xref="taxon:9606"
 175..3942
 /note="c-erb-B-2 precursor"
 /codon_start=1
 /protein_id="CAA27060.1"
 /db_xref="GI:31198"
 /db_xref="SWISS-PROT:P04626"
 /translation="MELAAALRWGLLLALLPBPAAASTQVCTGDMKRLRLPASPETHLD MLRLHYQCVQVQNLLELYPTNASLSFLQDIQEVGYVLIHNVQVPLQLRIY RGLTFEDNYALAVLDNGDPLNPTPTVGTASPGGLRELQLRLSLTEILKGGVLIQRNPQ LCYQDITLWKDIDFKHNNOLALTLDITNRSRACHPCSPMKSGRCWGESSEDCQSLTRT VCAGCARCKPDLNDCHEOCACAGCTGPKHSDCLACLHNSHCICELHCPALVYNT DTFESMPNPEGRYTFGASCVTACPNYLYSTVYGSCITLYCPLHNGEVTAEDSTQCEKC SKPCARVCIQGLMEHLREVRAVTSANIQEFAGCKIFGLAFPESEFDGDPASNTAPL QPEQLVFTEETIYLYISAWPDSLPLSFQNLQVIRPHQLLHNGAYSLTLQGLGI SACLRLSREGLSALHIIHNLHLCFVHTVPDQFLRNPHQALLHNRPEDECVGEG LAGLCARHCWGPGTQCYNCSQFLRGQCEVECVLQGLPREYVNAHCLPCHPE COPONGSVTCFPGPADOCVCAHYKDPFCVARGCPGVKPDLSYMPWKFPDEEGACO PCPINTHSCVDLDDKGCAPAEORASPLTSIIISAVVGILLVVLGVVFGILLKRRQKI RKYIMRRLQTELVEPLTPSGAMPNQAMRILKTELKRVKVLGSGAFIVYKGIWI PDGENVKIPAVKILRENTSPKANKELIDEAYVMAGVSPYVSRLLGICLTSTVOLVT QLMYPGCLLDHVRNRLGSLQDLNMCQIAKMSYLEDLVRLHROLAARNVLKSP NHVITDGLARLLDIDETEHADGKVPIKWMALSLRRRFTHQSDVWSYGVTVWE LMTFCARPDGIPAREIPDLLEKGERLPPOPICTIDVYIMVWKWMDISECRPRFREL VSEFSRMARDPQRFVIONEDLGPASPLDSTFYRSLLEDDDMGLDVAEEYLVPOQGF FCPDPAPGAGMVRHRRSSSTRSGGDLTLGLPSEEAEPRLAPSEAGSVDVFG DLGMAAGLSQSLTPHSDPQRYSESDTLPVLPSETDGYVAPLTCSPQPEYVNPQDVR POPPSRPFGLPAPADLDERPKTISPGKNVYKQVFAEGAVENPEYLTPOGGAA PQHPHPPAFPAFNLYWDQDPPERGPAPPSTFKGTPTAENPEYVLGDVPV"
 175..237
 /note="put. signal peptide (aa -21 to -1)"
 238..3939
 /product="put. c-erb-B-2 protein (aa 1-1234)"
 376..384
 /note="put. glycosylation site"
 544..558
 /note="put. glycosylation site"
 733..741
 /note="put. glycosylation site"
 949..957
 /note="put. glycosylation site"
 1762..1770
 /note="put. glycosylation site"
 1885..1893
 /note="put. glycosylation site"
 2059..2067
 /note="put. glycosylation site"

misc_feature
 /note="pot. glycosylation site"
 2353..3132
 /note="aa 727-986, seq. homologous to EGF receptor kinase domain"
 2446..2454
 /note="pot. glycosylation site"
 4455..4460
 /note="put. polyA signal"
 4473
 /note="polyA site"
 BASE COUNT 902 a 1383 c 1329 g 859 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.64e-53 Length: 4473
 Score: 1450.00 Matches: 266
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 SEQ4 (1-266) x HSERB2R (1-4473)
 QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
 |||||
 Db 3142 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
 |||||
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGly 40
 |||||
 Db 3202 GAGGACGATGATGGGGGAGCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3261
 |||||
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60
 |||||
 Db 3262 TTCTTCTGTCAGACCCTGCCCGGGCGCTGGGGCATGGTCCACACAGGACCCGACG 3321
 |||||
 QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGlyProSerGluGlu 80
 |||||
 Db 3322 TCATCTACGAGAGTGGCGGTGGGACCTGACACTAGGCTGGAGCCCTCTGAAGAGGAG 3381
 |||||
 QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
 |||||
 Db 3382 GCCCCACGGTCTCCACTGGCACCTCCGAAAGGGGCTGGCTCGGATGTATTTGATGGTGAC 3441
 |||||
 QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
 |||||
 Db 3442 CTGGGAATGGGGGAGCCAGAGGGCTGCAAGCCCTCCACACACATGACCCAGCCCTCTA 3501
 |||||
 QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
 |||||
 Db 3502 CAGCGGTACAGTGAGGACCCACAGTACCCTCGCCCTCTGAGACTGATGGCTACGTTGCC 3561
 |||||
 QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
 |||||
 Db 3562 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACACGACAGATGTTTCGGCCCCCAGCCC 3621
 |||||
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
 |||||
 Db 3622 CTTTGGCCCCAGAGGGGCCCTCTGCCTGCTGCCGACCTGCTGGTGGCCACTCTGGAAAGG 3681
 |||||
 QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
 |||||
 Db 3682 CCCAAGACTCTCTCCCCAGGAGAAGTGGGGTGGTCAAGAGCGTGTTCCTTTGGGGGGT 3741
 |||||
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
 |||||
 Db 3742 GCCGTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGGAGTGGCCCTTCAGCCCCACCCCT 3801
 |||||
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
 |||||
 Db 3802 CCTCCTGCCTTTCAGCCCCAGCCTTCGACAACTCTATTACTGGGACAGGAGCCACCCAGAG 3861
 |||||
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
 |||||

```

Db      3862  CGGGGGCTCCACCCAGCACCCTTCAAAGGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3921
Qy      261  GlyLeuAspValProVal 266
        |||||||
Db      3922  GGTCTGGACGTGCCAGTG 3939

RESULT 12
AX384608
LOCUS      AX384608                1806 bp      DNA      linear      PAT 19-MAR-2002
DEFINITION Sequence 5 from Patent WO0214503.
ACCESSION  AX384608
VERSION     AX384608.1  GI:19577809
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     McNeill,P.D. and Vedvick,T.S.
TITLE       Hand-Zimmermann,S., Cheever,M.A., Foy,T.M., Lodes,M.J., Kalos,M.D.,
            Compositions and methods for the therapy and diagnosis of
            her-2/neu-associated malignancies
JOURNAL     Patent: WO 0214503-A 5 21-FEB-2002;
            CORIXA CORPORATION (US)
FEATURES    source
            location/Qualifiers
            1..1806
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT 391 a 530 c 544 g 341 t
ORIGIN

Alignment Scores:
Pred. No.:      1.66e-53      Length:      1806
Score:          1442.00      Matches:    265
Percent Similarity: 99.62%      Conservative: 0
Best Local Similarity: 99.62%      Mismatches: 1
Query Match:     99.45%      Indels:    0
DB:              6          Gaps:      0

SEQ4 (1-266) x AX384608 (1-1806)

Qy      1  GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
        |||||||
Db      1003  CAGAATGAGACTTGGCCCGCAGCAGCTCCCTTGCAGCAGCACCTTCTACCGCTCACTGCTG 1062
Qy      21  GluAspAspMetGlyAspLeuValAlaGluGluTyrLeuValProGlnGlnGly 40
        |||||||
Db      1063  GAGGACGATCAGATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCCGCAGCAGGC 1122
Qy      41  PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
        |||||||
Db      1123  TTCTTCGTCCAGACCTGCCCGGGCGCTGGGGGATGCTCCACACAGCAGCAGCAGC 1182
Qy      61  SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
        |||||||
Db      1183  TCATCTACCAGAGTGGCGGTGGGACCTGACACTAGGGCTGAGGCCCTCTGAAGAGGAG 1242
Qy      81  AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
        |||||||
Db      1243  GCCCCAGGCTCTCACTGGCAGCCCTCGAAGGGGCTGGCTCCGATGATTTGATGGTGAC 1302
Qy      101  LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
        |||||||
Db      1303  CTGGGAATGGGGCAGCAAGGGGCTGCAAAAGCCTCCCCACACATGACCCCGCCCTCTA 1362
Qy      121  GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
        |||||||
Db      1363  CAGCGGTACAGTGAAGACCCACAGTACCCCTCCCTCTGAGACTGATGGCTACGTGCC 1422
Qy      141  ProLeuThrCysSerProGlnProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
        |||||||
Db      1423  CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGACAGCAGCATGTTGGCCCCCAGCCC 1482

```

```

Qy      161  ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
        |||||||
Db      1483  CCTCGCCCGCAGAGGGCCCTCTGCTTGCCTGCTGCCGACCTGCTGCTGCACCTCTGGAAAGG 1542
Qy      181  ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
        |||||||
Db      1543  GCCAAGACTCTCTCCCGAGGGAAGATGGGCTCTCAAAGACGTTTTTGGCTTTGGGGGT 1602
Qy      201  AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
        |||||||
Db      1603  GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGAGCTGCCCTCTACGCCCCACCT 1662
Qy      221  ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
        |||||||
Db      1663  CTTCTCTGCTTCAGCCCGCCTTCGACAACCTATTACTGGGACAGGACCCACACAGAG 1722
Qy      241  ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
        |||||||
Db      1723  CGGGGGCTCCACCCAGCAGCTTCAAAGGAGACCTACGGCAGAGAACCCAGAGTACCTG 1782
Qy      261  GlyLeuAspValProVal 266
        |||||||
Db      1783  GGTCTGGACGTGCCAGTG 1800

RESULT 13
AX505114
LOCUS      AX505114                3678 bp      DNA      linear      PAT 27-SEP-2002
DEFINITION Sequence 7 from Patent WO0240059.
ACCESSION  AX505114
VERSION     AX505114.1  GI:23386421
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Mincheff,M.S., Loukinov,D.I. and Zoubak,S.
TITLE       Methods and compositions for inducing cell-mediated immune
            responses
JOURNAL     Patent: WO 0240059-A 7 23-MAY-2002;
            American Foundation for Biological Research Inc. (US) ; Mincheff,
            Milcho S. (US) ; Loukinov, Dmitri I. (US) ; Zoubak, Serguei (US)
FEATURES    source
            location/Qualifiers
            1..3678
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            7..3678
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD48804.1"
            /db_xref="GI:23386422"
            /translation="MRLPASPTHLDMLRLHYOGCVQVGNLELYLPTNWSLSFLOD
            IQBVGQVLIAHNQVRQVPLRIYRGVLTQFDNVALVDNGDPLNNPTPYTGASP
            GLIRLEQLRSLEILKQGVLIQRPOLCYQDTILWKDIFHKNQLALTLDITNRSRAC
            HPCSPMKGRSCWGSESDCOSLTRVTCAGCARKGPIPTDCHEOCAGCTGPKHS
            DCLACLHNHNSICELHCPALVTYNTDTESMPNPRGYTFGASVCTACPYNLSTDV
            GSCTLVCPLNHOEVTAEDGTQCEKSKPCARVCYGLGMEHLREVRVTSANTQEFAG
            CKKIFGSLAFLPESFDGDPASNTAPLQPOVFETLEETITGLYLSAWPDSLPLSV
            FQNLQVIRKILHNGAYSUTLQGLGISWLGRLSRLGSLGLIHHNTHLUCFVHTVPM
            DOLFRPHQALLTANRDECEGELACHQCHGWCWGPCTQCVNCVSQFRLGQEC
            VESCRVQLGLPREYNARHCLPCPECPQNGSVTCFGEADQCVACAHKDPFCFVA
            RCPGVKPDLSYMPIWKFPDEGACOPCINCTHSCVDLDLDDKCPAORASPLTSIVS
            AVVGILLVVVLGVVFGILLKRRQOKIRKVTMRLLQETELVEPLTSGAMPNQAOMRI
            LKTELRLKVKLGSGAFYVYKGIWIPDGENVKIPVAIKVLRNTPSPKANKEILDEAY
            VMAGVSPYVSRLLGICLTSTVOLVTLMPYGCGLDLHDHVRNRRGLSQDLNNWQMLA
            KGMYSLEDVRLVRLHLLAARNLVKSPNHVKITDFGLARLIDIDITDYHADGGVPIKW
            MALESILRRRFTHQSDVSYGVTVVWELMTFGAKPYDGIPIAREIPLLEKEGRLPQPI
            CTIDVVMYWKWMDSECRPRELSEFSEARNARDPOREVIWINDLGLPASPLDSTF
            YRSLLEDDMDGLVDAEEVLPQOQFCFCDPDPAPCAGGMVHHRSSSTRSGSGDGLTLG
            LEPSERAPRSPPLASPEGAGSDVFDGLMGAAKGLQSLPTHDPSPLOQRYSEDPYPL
            PSTETDGYVAPLTCSPQPEYVNPDPVRPQPPSPREGPLPAARPATLTERAKTILSPGKN

```


GVKDVFAFGGAVENPEYLTTPQGAAPQHPHPAFSPAFDNLYWDQDPPPERGAPPST
FKGTPTENPEYLGIDVPV"
BASE COUNT 746 a 1137 c 1091 g 704 t
ORIGIN

Alignment Scores:
Pred. No.: 3.04e-53 Length: 3678
Score: 1442.00 Matches: 265
Percent Similarity: 99.62% Conservativeness: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.45% Indels: 0
DB: 6 Gaps: 0

SE04 (1-266) x AX505114 (1-3678)

```
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2878 CAGAAAGAGGACTTGGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 2937
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 2938 GAGGACGATGATGAGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 2997
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 2998 TTCTTCTCTCAGACCCCTGCCCGGGCTGGGGCATGGTCCACACAGCAGCCGCGCAGC 3057
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGluGluProSerGluGluGlu 80
Db 3058 TCATCTACCGAGGAGTGGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3117
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3118 GCCCCAGGCTCTCCACTGGCACCTTCCGAAAGGGCTGGCTCCGATGATTTGATGGTGAC 3177
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3178 CTGGGAATGGGGGAGCCAGGAGGGCTGCAAAAGCCTCCCCACACATGACCCCGAGCCCTCTA 3237
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3238 CAGCGTACAGTGAAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGGCTAGCTTGCC 3297
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3298 CCCCTGACCTGACGCCCCCAGCCTGAATATCTGAACAGCAGATGTTGGCCCCAGCCG 3357
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3358 CCTTCGCCCGGAGAGGGGCTCTGCTGCTGCCGACCTGCTGGTGCCACTCTGGAAGAG 3417
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3418 GCCAAGACTCTCTCCCGAGGAAGAATGGGGCTCGTCAAAAGCGTTTGGCTTTGGGGGT 3477
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3478 GCCGTGGAGAACCCGAGTACTTGACACCCCGAGGAGGAGTGGCCCTCAGCCCGACCCCT 3537
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3538 CCTCCTGCTTTCAGCCCGAGCTTCGACAACTCTATTACTGGGACAGGACCCAGCAGAG 3597
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3598 CGGGGGGCTCCACCCAGCACCTTCAAGGGGACCTACGGGAGAGAACCCAGAGTACCTG 3657
Qy 261 GlyLeuAspValProVal 266
Db 3658 GGTCTGGACGTGCCAGTG 3675
```

RESULT 14
AX060704

LOCUS AX060704 3768 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 2 from Patent WO0100244.
ACCESSION AX060704
VERSION AX060704.1 GI:12406101
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Erickson, S. and Schwall, R.
TITLE Methods of treatment using anti-erbB antibody-maytansinoid
conjugates
JOURNAL Patent: WO 0100244-A 2 04-JAN-2001;
Genentech, Inc. (US)
FEATURES
source Location/Qualifiers
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN

Alignment Scores:
Pred. No.: 3.1e-53 Length: 3768
Score: 1442.00 Matches: 265
Percent Similarity: 99.62% Conservativeness: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.45% Indels: 0
DB: 6 Gaps: 0

SE04 (1-266) x AX060704 (1-3768)

```
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAAGAGGACTTGGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGATGAGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTCTCTCAGACCCCTGCCCGGGCTGGGGCATGGTCCACACAGCAGCCGCGCAGC 3147
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGluGluProSerGluGluGlu 80
Db 3148 TCATCTACCGAGGAGTGGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3207
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCCCCAGGCTCTCCACTGGCACCCCTCCGAAAGGGCTGGCTCCGATGATTTGATGGTGAC 3267
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGAATGGGGGAGCCAGGAGGGCTGCAAAAGCCTCCCCACACATGACCCCGAGCCCTCTA 3327
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGTACAGTGAAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTAGCTTGCC 3387
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3388 CCCCTGACCTGCACCCCGGAGGAGGCTGCTGCCCTGCCGACCTGCTGGTGGCAGCCG 3447
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3448 CCTTCGCCCGGAGAGGCGCCCTCTGCCCTGCCGACCTGCTGGTGGCAGCTCTCGAAGAG 3507
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3508 GCCAAGACTCTCTCCCGAGGAAGAATGGGGTCTGCTCAAAAGACGTTTTTGGCTTTGGGGGT 3567
```


QY 201 AlavalGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Dy 3568 GCGGTGAGAACCCCGAGTACTTTCACACCCAGGAGGAGTCCCTCCAGCCCAACCT 3627
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTrpAspGlnAspProProGlu 240
Dy 3628 CTTCTGCTTACAGCCAGCTTCGACAACTCTATTACTTGGAGCAGGAGGAGGAGG 3687
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Dy 3688 CGGGGGGCTCCACCCAGCACCTTCAAGAGGACACCTAGGCGAGAGAACCCAGAGTACCTG 3747
QY 261 GlyLeuAspValProVal 266
Dy 3748 GGTCTGGACGTGCGAGTG 3765

RESULT 15
AX467229
LOCUS AX467229 3768 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0234287.
ACCESSION AX467229
VERSION AX467229.1 GI:21900511
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1
Beier, A.M., Gautam, A. and Mouritsen, S.R.
Novel therapeutic vaccine formulations
Patent: WO 0234287-A 3 02-MAY-2002;
Pharmexa A/S (DK)

FEATURES
source Location/Qualifiers
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD42452.1"
/db_xref="GI:21900512"
/translation="MELALCPWGLLLALLPCCAATVCTGTGDMKLRPLPASPETHLD
MRHLVQGVQVQGNLELYLPNASLSFLDIOEVGVVLIAHNOVQVPLQRLRV
RGQLFEDNTALAVLNDPLNNTPTVTPGASGLRELOLRSLTEILKGLVLRNPQ
LCYQDTILMKDIFPKXNQLALTIDTNRACHPCPMCKSRGWSESDCQSLVPT
VCAGGACRCKGPLPTDCECAAGCTGPKHSDCLACHFNHSGICELHCPALVYNT
DTESMPNPEGRVTFGACVACPYNLSTDVGSCVTLVGLHNOEVTAEQGTORCEK
SKPCARVCYGLGMEHLREVRVTSANIQFAGCKIFGSLAPLPESFDGDPASNTAPL
QPOLQVETLEITGLYXISAMPDSLPLSVQNLQVIRGLLHNGAYLSLQGLGI
SWGLSLRELGLSLALIHNTLHLCFVHTVPMQDLPRNPHQALLTANRPEDCEVGE
LACHQLCARGHCWPGPTQVNCQSLRQECVEECVILQGLPREYVNAHRLCPHPE
CQPNQSVTCFGEADQVACAHYKDPFCVAPCSGVKPDLSYMPVWFFGILIKRROKI
PCPINTHSCVDLDDKGCFAEORASPLTSIVSAVGLLVVGLVGVGILIKRROKI
RKYTMRLQETELVELPSPGAMPNOAMRILKETELBKVKVLSGAGTYVKGWI
PDGENVKIPVAKVLENTSPRANKELDEAVYVAGVSPVSRLLGICLTSVLQVLT
QLMPYGLLDHVRNGLSGDLNMQIAKMSYLEDLVLRHDLAARNVLRKSP
NHWKTDGLRLDLIDETEHADGKVPKIMMALESILRRFTHQSDVMSYGVTVWE
LMTFGAKPDGIPAREPDLLEKGERLPQPICTIDYIMVIMVKNMIDSECRPFREL
VSFESRMARDPQRFVIONDELGPASPLDSTFYRSLDDMDGLVDAYEYLVPQGF
FCPDAPGAGGVHHRSSSTSGGGDLTLGLEPSEEPAPSPAPSGAGSDVFDG
DLGMGAAGLQSLPHTDPFLQRYSEDPTVPLPSEDTGVYVPLTCSFQPEYVNPQDVR
PQPPREGPLPAARPAALERAKTLSPGKNGVKVDVFAFGVAGNPEYLTPOGGAA
PQHPHPAPSPAFDNLVYWDQPPGAPPPFTFKGTPTAENPEYLGILDVVPV"
BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN

Alignment Scores:
Pred. No.: 3,1e-53 Length: 3768
Score: 1442.00 Matches: 265
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.45% Indels: 0

Search completed: October 16, 2003, 11:03:03
Job time: 3853.81 secs

DB: 6 Gaps: 0
SEQ4 (1-266) x AX467229 (1-3768)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Dy 2968 CAGAAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACAGCACCTTCTACCCGCTCACTGCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Dy 3028 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGAGATATCTGGTACCCAGCAGGCG 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60
Dy 3088 TTCTTCTGTCCAGACCTGTCGCCGGCGCTGGGGGATGCTGCCACACAGCAGCAGCCGAGC 3147
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Dy 3148 TCATCTACCAAGGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3207
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Dy 3208 GCCCCAGAGTCTCCACTGGCACCTCCGAAAGGGCTGGCTCCGATGTATTTGATGCTGAC 3267
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisaspProSerProLeu 120
Dy 3268 CTGGGAATGGGGGAGCAGCCAAAGGGCTGCCAAAGGCTCCCCACACATGAGCCAGCCCTCTA 3327
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Dy 3328 CAGCGGTACAGTACGAGGACCCACAGTACCCCTGCTGAGACTGATGCTGCTGCTGCTGCTG 3387
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Dy 3388 CCCCTGACCTGACGCCGCCAGCTGAATATGTGAACAGCAGATGTTCCGGCCCGCAGCCC 3447
QY 161 ProSerProArgGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Dy 3448 CCTTCCCGCCGAGAGGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3507
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Dy 3508 GCCAAGACTCTCTCCCGCAGGAGGAGATGGGCTGCTCAAGACCGTTTGTGCTTTGGGGGT 3567
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Dy 3568 GCGGTGAGAACCCCGAGTACTTTCACACCCAGGAGGAGTCCCTCCAGCCCAACCT 3627
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTrpAspGlnAspProProGlu 240
Dy 3628 CTTCTGCTTACAGCCAGCTTCGACAACTCTATTACTTGGAGCAGGAGGAGGAGGAGG 3687
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Dy 3688 CGGGGGGCTCCACCCAGCACCTTCAAGAGGACACCTAGGCGAGAGAACCCAGAGTACCTG 3747
QY 261 GlyLeuAspValProVal 266
Dy 3748 GGTCTGGACGTGCGAGTG 3765

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 5110.68 Seconds
(without alignments)
3105.423 Million cell updates/sec

Title: SEQ3
Perfect score: 3628
Sequence: 1 MELAALCRWGLLLALLPPGA.....CVDLDDKCPAEQRASPLTS 653

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODEL=frame_plus_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool_P/HOLLERAN480/runat_15102003_131913_20548/app_query.fasta_1.4685
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGN_1_1_13645 -runat_15102003_131913_20548 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hct.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hct.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_mam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_fod.*
- 26: em_gss_phg.*
- 27: em_gss_vrl.*
- 28: gb_gssl.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3087.5	85.1	4323	11	AK031099	AK031099 Mus muscu
2	3087.5	85.1	4463	11	AK083669	AK083669 Mus muscu
3	1533.5	42.3	2456	11	AK004911	AK004911 Mus muscu
4	1533.5	42.3	2662	11	AK004883	AK004883 Mus muscu
5	1533.5	42.3	2936	11	AK004944	AK004944 Mus muscu
6	1483	40.9	964	13	BUI50809	BUI50809 AGENCOURT
7	1410	38.9	1016	12	BM562913	BM562913 AGENCOURT
8	1364	37.6	757	9	AU140362	AU140362 AU140362
9	1339	36.9	3110	11	AK031542	AK031542 Mus muscu
10	1333	36.7	808	14	CA489534	CA489534 AGENCOURT
11	1293	35.6	894	10	BE746725	BE746725 601579159
12	1205.5	33.2	795	14	CA328613	CA328613 UI-M-FY0-
13	1194.5	32.9	1201	13	AX402419	AX402419 BX402419
14	1140	31.4	614	9	AW370693	AW370693 QV1-BT026
15	1137	31.3	613	9	AW410534	AW410534 fh06h06.x
16	1126	31.0	659	9	AW057736	AW057736 wx02h09.x
17	1115.5	30.7	1050	10	BF984645	BF984645 602309702
c 18	1097	30.2	670	9	AI906012	AI906012 RC-BT105-
19	1094	30.2	733	12	BI115219	BI115219 602861672
20	1073	29.6	588	2	HSN068717	Bx478931 Homo sapi
21	1063	29.3	563	10	BE297998	BE297998 601118462
22	1049	28.9	583	2	HSN068900	Bx479114 Homo sapi
23	1046	28.8	569	9	AL701765	AL701765 DKF2p686p
c 24	933	25.7	669	10	BF108852	BF108852 7164h02.x
25	915	25.2	641	10	BB621058	BB621058 BB621058
c 26	914	25.2	872	13	BX391706	BX391706 BX391706
27	893	24.6	474	9	AL701569	AL701569 DKF2p686A
28	880	24.3	561	12	BM720098	BM720098 UI-E-E00-
29	869	24.0	876	13	BQ768889	BQ768889 UI-M-F10-
c 30	847	23.3	514	10	BF998814	BF998814 QV0-GN014
31	842.5	23.2	657	10	BB620332	BB620332 BB620332
32	839.5	23.1	632	10	BB655484	BB655484 BB655484
33	830	22.9	539	14	CA544382	CA544382 C0648D06-
34	811	22.4	923	13	BQ888953	BQ888953 AGENCOURT
c 35	788.5	21.7	499	10	BF746234	BF746234 RCL1-BT025
c 36	788.5	21.7	499	10	BF746304	BF746304 RCL1-BT025
c 37	761	21.0	1057	12	BM907440	BM907440 AGENCOURT
c 38	747.5	20.6	486	9	AI906364	AI906364 RC-BT108-
39	743	20.5	447	10	BF436479	BF436479 7p13c10.x
c 40	738	20.3	995	13	BX336779	BX336779 BX336779
c 41	734	20.2	1201	13	EX402418	EX402418 BX402418
42	729	20.1	450	9	AI906050	AI906050 RC-BT103-
43	729	20.1	508	10	BE065758	BE065758 RC2-BT031
44	723.5	19.9	660	12	BI557977	BI557977 603236665
c 45	709	19.5	455	12	BG991986	BG991986 MR2-HT116

ALIGNMENTS

RESULT 1
AK031099
LOCUS AK031099 4323 bp mRNA linear HTC 05-DEC-2002
DEFINITION Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length
enriched library, clone:5930404N10 product:v-erb-b2 erythroblastic
leukemia viral oncogene homolog 2, neuro/glioblastoma derived
oncogene homolog (avian), full insert sequence.
ACCESSION AK031099
VERSION AK031099.1 GI:26082143
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AK031099

Db 468 CTGCAGCTCGAAGTCTCACAGATCTTGAAGGAGGAGTGTTCATCCGTGGGAACCT 527
 Qy 160 GlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeu 179
 Db 528 CAGCTCTGCTACCAAGAGATGTTTGTGGAGGATGTCCTCCGTAAGAAATACACAGCTG 587
 Qy 180 AlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCys 199
 Db 588 GCTCCTGTCGACATGCACCAATCGTTCCTGGGCTGTCCACCTGTGCCCCAACCTGC 647
 Qy 200 LysGlySerArgCysTrpGluSerSerGluAspCysGlnSerLeuThrArgThrVal 219
 Db 648 AAGACAATCACTGTGTGGGTGAGAGTCTCTGAAGACTGTCAAGTCTTGCATGGCACCATC 707
 Qy 220 CysAlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGln 239
 Db 708 TGTACTAGTGTGCTGCGCGGTGCAAGGCGCGCTGCCCACCTGACTGTGTGCATGAGCAG 767
 Qy 240 CysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsn 259
 Db 768 TGTGCTGCAGCTGCACGGTCCCAAGCATTTCTGACTGACTGGCTGCTCCACTTCAAT 827
 Qy 260 HisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPhe 279
 Db 828 CATAGTGTATCTGTAGCTGACTGCGCGCGCTCATCACTCACTACACACAGACACCTTC 887
 Qy 280 GluSerMetProAsnProGluCysLysArgThrPheGlyAlaSerCysValThrAlaCys 299
 Db 888 GAGTCCATCTCAACCTCGAGGGTGCCTACACCTTTGTTGTCAGCTGTGACCACTGC 947
 Qy 300 ProTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsn 319
 Db 948 CCTTACAACCTACCTCTCCACGAAAGTGGATCTCTGCTGCTGCTCCCGCAACAC 1007
 Qy 320 GlnGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAla 339
 Db 1008 CAAGAGGTTCACAGCTGAGGACGGAACACAGCGGTGTGAGAAATGAGCAAGCCCTGTGCT 1067
 Qy 340 ArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAla 359
 Db 1068 GGAGTATGCTATGCTGCGGATGGAGCACCTCCGAGGGGCGAGGCCATCACCAGTGAC 1127
 Qy 360 AsnIleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGlu 379
 Db 1128 AATATCCAGGAGTTGCTGCTGCAAGAAGATCTTGGGAGCTTGGCATTTTCCGCGAG 1187
 Qy 380 SerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnVal 399
 Db 1188 AGCTTTGATGGAAACCCCTCTCCGCGGTTCGCCACTGTAAGCCAGAGCATCTCCAAGTG 1247
 Qy 400 PheGluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeu 419
 Db 1248 TTGGAAACCTTGGAGGAGATCACAGTTTACTATATCATTTCCAGATGCCAGAGAGCTTC 1307
 Qy 420 ProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGly 439
 Db 1308 CAAGACCTCACTGCTTCCACAACCTTCGGGTTCATTCGGGCGAGGATTTCTCCATGATGT 1367
 Qy 440 AlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArg 459
 Db 1368 GCTTACTCATTTGACGTTTGAAGGCTGGGATTCACCTCACTPCTGGGCTACGCTCACTGCGG 1427
 Qy 460 GluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThr 479
 Db 1428 GAGCTGGCACTGATGGCTCTCATTCACCGCAACACCCATCTCTGCTTTGTAACACT 1487
 Qy 480 ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg 499
 Db 1488 GTACCTTGGGACCACTCTTCCGGAACCCGACACGAGCCCTACTCCAGTGGGAACCGG 1547
 Qy 500 ProGluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHis 519

Db 1548 CCAGAAGAGCATGTGTGCTTTCAGGGCTTGGTCTCTGTAACCTCACTGTGTGCCGTGGGCAC 1607
 Qy 520 CysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGlu 539
 Db 1608 TGCTGGGGGCCAGGGCCACCAGTGTCTAACTGCAGTTCAGTTCTCCGGGGCCAGGAG 1667
 Qy 540 CysValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHis 559
 Db 1668 TGTGTGGAGGAGTCCGAGTATGGAAGGGCTCCCGAGGAGTATGTGAGGGGCAAGCAG 1727
 Qy 560 CysLeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyPro 579
 Db 1728 TGTCTGCCATGCCACCGAGTGTGAGGCTCAAAACAGCTCGGAGACCTGCTATGATCG 1787
 Qy 580 GluAlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArg 599
 Db 1788 GAGGCTGACCACTGTGAGGCTGTGCCACACTACAAGGACTCATCTTCCTGTGTGCTCGC 1847
 Qy 600 CysProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGlu 619
 Db 1848 TGCCCCAGTGTGTGAAGCCAGACCTCTCTACATGCCTATCTGGAAGTACCCGATGAG 1907
 Qy 620 GluGlyAlaCysGlnProCysProLysAsnCysThrHisSerCysValAspLeuAspAsp 639
 Db 1908 GAGGCGATATGTCAGCCATGCCCATCACTCACTGCACCCACTCATGTGTGGACCTGGACGAA 1967
 Qy 640 LysGlyCysProAlaGluGlnArgAlaSerProLeuThr 652
 Db 1968 CGAGCTGCCAGCAGAGCAGAGCCAGCCAGCTGAGCA 2006

RESULT 2
 AK083669
 LOCUS
 DEFINITION
 Mus musculus 9 days embryo whole body cDNA, RIKEN full-length enriched library, clone:D030063B12 product:v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian), full insert sequence.
 AK083669
 ACCESSION
 VERSION
 AK083669.1 GI:26101404
 HTCC; CAP trapper.
 KEYWORDS
 SOURCE
 Mus musculus
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
 2
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 MEDLINE
 PUBMED
 11042159
 REFERENCE
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 MEDLINE
 PUBMED
 11076861
 REFERENCE
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toyama,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

TITLE JOURNAL

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGCGCCGCACTCGAGTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence[5' GAGAGAGAGAGCGCCGCACTCAATTAATTAATTAACCCCGCCCGCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOUR.

COMMENT

FEATURES

source

Location/Qualifiers

1..2456
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:1300005M11"
/db_xref="MGI:1895040"
/db_xref="taxon:10090"
/clone="1300005M11"
/sex="male"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
256..2223
/note="unnamed protein product; epidermal growth factor receptor (MGD|MGI:95294)
putative"

CDS

/codon_start=1
/protein_id="BAB23662.1"
/db_xref="GI:12836453"
/db_xref="MGI:95294"
/translation="MRPSTARTLLVLTALCAAGALEEKVCOGTSNRLTQLTGF
EDHFLSLQRYNCEVVLGNLEITYVORNYDLSFLTKIQEVAGVLTALNTVERIPLE
NLQIIRGNALYENTYALAILNSYNTGRNLPRLNQLIELIGAVRFSNNPLICND
TIQWRDQVNFVMSMDLQSPKCDPNSCWGGGEECNCKLTKICAQ
CSHRGRSPSDCHNOCAAGCTGPRESCLVQKQDEATCKDTCPLMLNPTTYQ
MDYNEKYSFEGATCKVCKPNVYVDHSGVACGPDYVEVEEDGIRKCKKCDGPCR
KVCNGIGKEFKDTLSINATNIKFKYCTAISGLDLHLPLVAFKDSFTPTPLDREL
ELIKTVKEITGFLIQAQNPWDHLAFNELIKRTKRGQFSLAVGKLNITSLGL
LSKEISDGVIIISGRNLCYANTINMKLFGTPNOKTKIMNRBCKDKAVNVHCP
LCSECGWPEPDRCQVNSRGRECNLEGEPRFVENSICIQCHPECLPQA
MNTCTGGRPDNQCIAHYIDGPHCVKTCFAGTNGENNTLVWRYADANNVCHLCHANC
TYCAGPGLQGCBEVPSGYVQWQWILKTFWI"

BASE COUNT 631 a 670 c 620 g 535 t
ORIGIN

Alignment Scores:

Pred. No.: 3,98e-115 Length: 2456
Score: 1533.50 Matches: 284
Percent Similarity: 59.94% Conservative: 99
Best Local Similarity: 44.44% Mismatches: 233
Query Match: 42.27% Indels: 23

DB: 11 Gaps: 7
SEQ3 (1-653) x AK004911 (1-2456)
QY 11 LeuLeuLeuAlaLeuLeuProProGlyAlaAla-----SerThrGlnValCysThrGly 28
D 295 TTGCTGACCGCGCTCTCGCGCGAGGTGGCGCTTGGAGGAAAGAAAGTCTGCCAAGGC 354
QY 29 ThrAspMetLysLeuArgLeuProAlaSerProGluThrHisLeuAspMetLeuArgHis 48
D 355 ACAAGTAACAGGCTCACCCAACTGGGCATCTTTGAGAGACCCTTTCTGAGCCCTGCAGAGG 414
QY 49 LeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuLeuLeuTyrTyrLeuProThrAsn 68
D 415 ATCTACAACAACCTGTGAAGTGGTCTTGGGAACCTTGGAATTTACCTATGTGCAAGGAAT 474
QY 69 AlaSerLeuSerPheLeuGlnGlyGlnGluValGlnGlyTyrValLeuLeuAlaHis 88
D 475 TAGGACCTTTCTTCTTAAAGACCATCCAGAGGTGGCGGCTATGTCTCATTTGCCCTC 534
QY 89 AsnGlnValArgGlnValProLeuGlnArgLeuArgTleValArgGlyThrGlnLeuPhe 108
D 535 AACACCTGGAGAGATCCCTTTGGGAACCTGCAGATCATCAGGGGAATATGCTCTTAT 594
QY 109 GluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspProLeuAsnAsnThrThrPro 128
D 595 GAAACACCTATGCCITAGCCATCTTCCCAAC----- 627
QY 129 ValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeuArgSerLeuThrGluLe 148
D 628 ---TATGGGACAACAAGAACTGGGCTTAGGGAATGCCCATGGGAATTTACAGGAATC 684
QY 149 LeuLysGlyGlyValLeuLeuGlnArgAsnProGlnLeuCystyrGlnAspThrIleLeu 168
D 685 CTGATTTGGTCTGTGGATTCAGCAACAACCCATCTCTGCAATATGATGATATCCAG 744
QY 169 TrpLysAspIle-----PheHisLysAsnAsnGlnLeuAlaLeuThrLeuLe 184
D 745 TGGAGGACATCTGTCACAAACGCTTTTATGAGCAACATGTCATGGACTTA----- 795
QY 185 AspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySerArgCys 204
D 796 ---CAGAGCCATCCGACAGTTCGCCCAATGTGATCAAGCTGTCCCAATGGAAGTGC 852
QY 205 TrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAlaGlyGlyCys 224
D 853 TGGGGAGGAGGAGGAGAACTGCCAGAAATTCACCAAAATCATCTGTGCCCAATGT 912
QY 225 Ala---ArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAlaAlaGly 243
D 913 TCCCATCGCTGCTGGCAGGTCCCGAGTGTGCTGCCAACCAACCAATGTGCTGCGGG 972
QY 244 CysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGlyLe 263
D 973 TGTACAGGCGCCGAGAGAGTGTCTGTCTGCCAAAGTTCCAAGATGAGGCCACA 1032
QY 264 CysGluLeuHisCysProAlaLeuLysThrTyrAsnThrAspThrPheGluSerMetPro 283
D 1033 TGAAGGACACCTGCCCAACCATCATGCTGTACACCCACCATCATCATGATGATGATG 1092
QY 284 AsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsnTyr 303
D 1093 AACCCCTGAAGGGAAGTACAGCTTTGTGTCACCTGTGTGAAGAAGTGTGCCCGAACTAC 1152
QY 304 LeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnValThr 323
D 1153 GTGGTGACAGATCATGGCTCATGTGTCGAGCGCTGTGGCGCTGACTACTACGAGAGT--- 1209
QY 324 AlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaAlaGlyCysTyr 343
D 1210 GAAGAGATGCGATCCGCAAGTGTAAAAAATGTGATGGCGCTGTCCCAAGTTGTAAT 1269
QY 344 GlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsnIleGlnGlu 363

Db	1476	ANGGAATTAACAGGCTTTTTCCTGATGTCAGGCTGGCCCTGGATGAATAACTGGACTGACCTCCAT	1535
Qy	424	ValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAlaIleSerLeu	443
Db	1536	GCTTTCGAGAACCTAGAAATATACTGTCGACAAACCAACATGCTCAGTTCTTCTTG	1595
Qy	444	ThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlyLeuGlySer	463
Db	1596	GGGGTCTGTTGGCCTGAACATCACTACATCGTGGGCTGCGTTCCCTCAAGGAGATCAGTGAT	1655
Qy	464	GlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrValProTrpAsp	483
Db	1656	GGGGATGTCATCTCTCGGAACCGAATTTGCTACGCAACAAACAATAAATCGAAA	1715
Qy	484	GlnLeuPheArgAsnProHisGlnAlaLeuHisThrAlaAsnArgProGluAspGlu	503
Db	1716	AAACTCTTCGGGACACCCCAATCAGAAAACCAAAATCATGAACACAGAGCTGAGAAGAC	1775
Qy	504	CysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTrpGlyPro	523
Db	1776	TGCAAGGCGGTGAACCCAGCTGTCGAATCCTTTATGCTCCTCGGAAGGCTGCTGGGGCCCT	1835
Qy	524	GlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCysValGluGlu	543
Db	1836	GAGCCAGGACTGTGCTCTCCGCCAAGATGTGACAGAGCGAGGAGTGGTGGAGAAA	1895
Qy	544	CysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCysLeuProCys	563
Db	1896	TGCAACATCTGGAGGGGGAACCAAGGAGTTTGTGGNAATTTCTGAATGCATCCAGTGC	1955
Qy	564	HisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGluAlaAspGln	583
Db	1956	CATCCAGATGTCTGCCCGCCATGAACATCACCTGTACAGCGAGCGGACACGACAAC	2015
Qy	584	CysValAlaCysAlaHisTyrIlysnProProPheCysValAlaArgCysProSerGly	603
Db	2016	TGATCCAGTGTGCCCACTACATTGGCCACACTGTGTCAAGACCTGCCAGCTGGC	2075
Qy	604	ValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlyAlaCys	623
Db	2076	ATCATGGGAGACAACACACTGTG--GTCGGAAAGTATGCAGATGCCAATAATGTCGC	2132
Qy	624	GlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLysGlyCys	642
Db	2133	CACCTATGCCAGCCCACTGATACCTATGGATGTCTGGCCAGGTCTTCAAGGATGT	2189
RESULT	5		
AK004944			
LOCUS			
DEFINITION			
AK004944		2936 bp mRNA linear	HTC 05-DEC-2002
VERSION		Mus musculus adult male liver cDNA, RIKEN full-length enriched	
KEYWORDS		library, clone:1300008123 product:epidermal growth factor receptor,	
SOURCE		full insert sequence.	
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBLISHED			
COMMENT			


```

QY 464 GlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrValProTrpAsp 483
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1649 GGGGATGTGATCATTTCTGGAACCGAATTTGTCTACGCAACACAACTAACTGGAA 1708

QY 484 GlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGluAspGlu 503
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1709 AAATCTCTCGGGACACCAACCAATCAGAAACCAACCAAAATCATGAACAGAGCTGAGAAGAC 1768

QY 504 CysValGluGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTrpGlyPro 523
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1769 TCAAGCGCGTGAACACGCTGCAATCTTTATGCTCTCGAAGGCTGTGGGGCCCT 1828

QY 524 GlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCysValGluGlu 543
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1829 GAGCCAGGGAGCTGTCTCTGCCAGAAATGTGACAGAGCGAGGAGTGGTGAGAA 1888

QY 544 CysArgValLeuGlnGlyLeuProArgGluTrpValAsnAlaArgHisCysLeuProCys 563
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1889 TGCAACATCTCTGGAGGGGAACCAAGGAGTTTGTGAAATTTCTGAATGTCATCCAGTGC 1948

QY 564 HisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGluAlaAspGln 583
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1949 CATCAGAAATGTCTCCCGAGCCCATGAACATCACTGTACAGGCGAGGGACCAAC 2008

QY 584 CysValAlaCysAlaHisTrpLysAspProPheCysValAlaArgCysProSerGly 603
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2009 TGCATCCAGTGTGCCCACTACATTCATGCCCCACACTGTGTCAAGACCTGCCACCTGGC 2068

QY 604 ValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGluGlyAlaCys 623
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2069 ATCATGGGAGGACAACTCTG---GTCTGGAAGTAGTGCAGATGCCAATAATCTCTGC 2125

QY 624 GlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLysGlyCys 642
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2126 CACCTATGCCAGCCCACTGACCTATGATGTGTGGGCCAGGCTTCAAGGATGT 2182

RESULT 6
LOCUS BUI50809
DEFINITION AGENCOURT_8118591 Lupski_dorsal_root_ganglion Homo sapiens CDNA
clone IMAGE:6180101 5', mRNA sequence.
ACCESSION BUI50809
VERSION BUI50809.1 GI:22664341
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 964)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13562 row: f column: 06
High quality sequence start: 27
High quality sequence stop: 529.
Location/Qualifiers
1. .964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6180101"
/sex="male"
/tissue_type="dorsal root ganglia"
FEATURES
source

```

```

/dev_stage="adult, 36 yr"
/lab_host="DH108"
/clone_lib="Lupski_dorsal_root_ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGACCCACGCGCTCG-3' and
5'-GACTAGTCTAGATCGGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

```

BASE COUNT 190 a 296 c 291 g 185 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 1.41e-111 Length: 964
Score: 1483.00 Matches: 287
Percent Similarity: 93.83% Conservative: 2
Best Local Similarity: 93.18% Mismatches: 12
Query Match: 40.88% Indels: 8
DB: 13 Gaps: 2

SEQ3 (1-653) x BUI50809 (1-964)

```

QY 319 AsnGlnGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCys 338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 AACCAAGAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGT 92

QY 339 AlaArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSer 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 93 GCCCAGAGTGTGATGGTGTGGCATGGAGCACTTGCAGAGAGTGGAGGCGACTTACCACT 152

QY 359 AlaAsnIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuPro 378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 153 GCCAATATCCAGGAGTGTGCTGGCTGCAAGAAGATCTTTGGGAGGCTTGCATTTCTGCCG 212

QY 379 GluSerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGln 398
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 213 GAGAGCTTTGATGGGACCCAGCTTCCAACTGCCCCCTCCAGCCAGAGCAGCTCCAA 272

QY 399 ValPheGluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSer 418
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 273 GTGTTGAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCGGACAGC 332

QY 419 LeuProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsn 438
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 333 CTGCCCTGACCTCAGCGTCTTCCAGAACCTTGAAGTAATCCGGGGACGAATTTGCAACAAT 392

QY 439 GlyAlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeu 458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 393 GCGCGCTACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCAGTG 452

QY 459 ArgGluLeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHis 478
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 453 AGGGAACCTGGGCACTGGACTGGCCCTCATCCACCATAACACCCACTCTCTGCTCGTGAC 512

QY 479 ThrValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsn 498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 513 ACGGTGCCCTGNGACCACTCTTTTCGGAACCCGCAACCAAGCTCTCTCCACACTCCCAAC 572

QY 499 ArgProGluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArg-G1 518
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 573 CGGCCAGAGGAGAGTGTGTGGCGAGGGCCCTGGCTGCCACCACTGTGGCCCGCAGGG 632

QY 518 yHisCysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyG1 538
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 633 GCATGCTGGGGTCCAGGGCCCAACCCAGTGTGTCAACTGCAGCCAGTTCCTTTCGGGGCCA 692

QY 538 nGluCysValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArg 558
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 693 GGAGTGGCTGGAGGAATCCGAGTACTGCAGGGGCT-CCCAGGGAGTATGTGAATGCCCA 751

```


COMMENT

Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel.: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES

source

Location/Qualifiers
1..757

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE2000402"
/tissue_type="placenta"
/clone_lib="PLACE2"
/note="Vector: pME18SFL3"

BASE COUNT 149 a 227 c 231 g 145 t 5 others

ORIGIN

Alignment Scores:

Pred. No.: 5,88e-102 Length: 757
Score: 1364.00 Matches: 246
Percent Similarity: 98.40% Conservative: 0
Best Local Similarity: 98.40% Mismatches: 4
Query Match: 97.60% Indels: 0
DB: 9 Gaps: 0

SEQ3 (1-653) x A0140362 (1-757)

QY 332 GluLysCysSerLysProCysAlaArgValCysTyrGlyLeuGlyMetGluHisLeuArg 351
DB 3 GAGAAGTCGACGAAGCCCTGCGCGAGTGCTATGCTGTGGGATGGAGCATTGGGA 62
QY 352 GluValArgAlaValThrSerAlaAsnIleGlnGluPheAlaGlyCysLysLysIlePhe 371
DB 63 GAGGTGAGGGCAGTTACCAGTGCCAAATATCCAGGAGTTGTGCTGCGCAAGAGATCTTT 122
QY 372 GlySerLeuAlaPheLeuProGluSerPheAspGlyAspProAlaSerAsnThrAlaPro 391
DB 123 GGGAGCCCTGGCATTTCTCCGGAGAGCTTGTATGGGGACCCAGCGCTCCAAACATCCGCCG 182
QY 392 LeuGlnProGluGlnLeuGlnValPheGluThrLeuGluGluIleThrGlyTyrLeuTyr 411
DB 183 CTCACGCCAGAGCAGCTCCAAGTGTTCAGACTCTGGAGAGATCACAGTTACCTATAC 242
QY 412 IleSerAlaTrpProAspSerLeuProAspLeuSerValPheGlnAsnLeuGlnValIle 431
DB 243 ATCTCAGCATGGCGGACAGCGCTGCTGACCTCAGCGTCTCCAGAACCTGCAAGTAATC 302
QY 432 ArgGlyArgIleLeuHisAsnGlyAlaTyrSerLeuThrLeuGlnGlyLeuSer 451
DB 303 CGGGAGCAATTCGCAATGGCCCTACTCGTGTACCTGCAAGGCTGGGCATCAGC 362
QY 452 TrpLeuGlyLeuArgSerLeuArgGluLeuGlySerGlyLeuAlaLeuIleHisAsn 471
DB 363 TGGCTGGGGCTGCGCTCACTGAGGAACCTGGGACGTGGACTGGCCCTCATCCACCATAAC 422
QY 472 ThrHisLeuCysPheValHisThrValProTrpAspGlnLeuPheArgAsnProHisGln 491
DB 423 ACCCACTCTGCTGTCGACACAGCGTGGCCGACCACTCTTTCGGAACCCGACCAA 482
QY 492 AlaLeuLeuHisThrAlaAsnArgProGluAspGluCysValGlyGluLeuAlaCys 511
DB 483 GCTCTGCTCCACACTGCCAACCGCCAGAGACGAGTGTGTGGGAGGGGCTGGCCTGC 542
QY 512 HisGlnLeuCysAlaArgGlyHisCysTrpGlyProGlyProThrGlnCysValAsnCys 531
DB 543 CACCAGCTGTGGCCCGGGGACACTGCTGGGGTCCAGGGCCACCACCCAGTGTGTCAACTGC 602

QY 532 SerGlnPheLeuArgGlyGlnGluCysValGluGluCysArgValLeuGlnGlyLeuPro 551
DB 603 AGCCAGTTCTTCGGGGCCAGGANTGCGTGGAGGAATCCGAGTACTGCANGGCTCCCC 662
QY 552 ArgGluTyrValAsnAlaArgHisCysLeuProCysHisProGluCysGlnProGlnAsn 571
DB 663 AGGAGTAGTGAATGCCAGGACACTGTTTGGCGTCCACCCCTGAGTGTGCAGCCCGAGAAT 722
QY 572 GlySerValThrCysPheGlyProGluAla 581
DB 723 GGCTCANTGACTGTTTGGACCGGANGCT 752

RESULT 9
AK031542
LOCUS
DEFINITION
Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
enriched library, clone:6030449F08 product:v-erb-b2 erythroblastic
leukemia viral oncogene homolog 2, neuro/glioblastoma derived
oncogene homolog (avian), full insert sequence.
AK031542
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
HTC: CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitzunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bul, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
and Hayashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851
 5

TITLE
JOURNAL
REFERENCE
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 3110)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukushima, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, Y., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL

Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.go.jp/
 URL:http://fantom.gsc.riken.go.jp/.

FEATURES
Source

Location/Qualifiers
 1..3110
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM.DB:6030449F08"
 /db_xref="taxon:10090"
 /clone="6030449F08"
 /sex="male"
 /tissue_type="testis"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="13 days embryo"
 <1..2648
 /note="unnamed protein product; putative v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian) (MGDI|GI:95410, GB|U71126, evidence: BLASTN, 99%, match=449)"
 /codon_start=3
 /protein_id="BAC27442.1"
 /db_xref="GI:26327397"
 /translation="AFLPESDPGNSGVPKPEHLQVFEETIGTYLTSAMPES FDTVQFLNRVIRGLHDGAYSLTLOGLTHSLRLSDLSGLAIRHNTLHCF VNTVPDQFLRNPQALLHSGNRPEACLEGVLCNSLCARGCWGPCTQVCNCSQF LRGOECVCECRWKGLPREYVRGKHLCPCHPCQONSETCYGSEADQCEACAHYKD SSSVCARCPGQPDLSPTWIKYIDEEGICQPCINPCHSCVDLDERGCPAEQRASPVTFIATVVGVLFIIVVIGILIKRRRKQIKYTMRLLETQELVEPSTPSGAVPN ILDEYVMAQVGSPTVSLGSGAFYKGIWIPDGNVAGIPVAIKVIREHRLGSDQLN IADYRVMAGVSGKLTSTVOLTPQMLPVGCLLDHVRHRLGSDQLN KVCVQIAKSLYEELVHRDLAARNVLYKSPNHVKINDPFLGLRLIDETEHADGG KYP1IKWMALESILRRRFTQSDWSYGVTVWELMTFGAPYDGFIPAREIPDLLEKGER LPQPICITDVTWIMVKMDSECRPRELVSFSESRNARDQRFVVIQNEIDLGPSS PMDSTVILSLEDDEDDMGAEGLVQDFSPDPAIGTGTSHRRHSRSRSGG GELTGLGPESEPRSPILAPSEVDGDLAVGVTIKGLQSLSPHDLSPQRYSE

CDS

[illegible]

RESULT 10

```

CA489534
LOCUS       AGENCOURT_10810594 MAPcL Homo sapiens cDNA clone IMAGE:6722008 5',
            mRNA sequence.
ACCESSION   CA489534
VERSION     CA489534.1 GI:24952325
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   NIH-MGC http://mgc.nci.nih.gov/.
            1 (bases 1 to 808)
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished
JOURNAL
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tissue Procurement: Kristi A. Eglund, Ira Pastan
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM14284 row: a column: 16
            High quality sequence stop: 582.
            Location/Qualifiers
                1..808
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6722008"
                /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1
                , LNCap"
                /lab_host="EMDH10B"
                /clone_lib="MAPcL"
                /note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
                Subtracted with brain, liver, lung, kidney and muscle.
                Directionally cloned. Priming method: oligo-dT. Average
                insert size: 1800 bp. Library amplification: 26,000 fold.
                Kristi A. Eglund, James J. Vincent, Robert Strausberg,
                Bungkok Lee & Ira Pastan: Discovery of new breast
                cancer genes encoding membrane and secreted proteins.
                Manuscript submitted."
BASE COUNT  169 a 255 c 225 g 159 t
ORIGIN
Alignment Scores:
Pred. No.:      2.28e-99      Length:      808
Score:          1333.00      Matches:    247
Percent Similarity: 94.80%      Conservative: 8
Best Local Similarity: 91.82%      Mismatches: 8
Query Match:     36.74%      Indels:     6
DB:              14          Gaps:       4

SEQ3 (1-653) x CA489534 (1-808)

QY 133 SerProGlyGlyLeuArgGluLeuGlnLeuArgSerLeuThrGluLeuLeuLysGly 152
    |||||
Db 3 TCCCCAGGAGGCGCTCGGGAGCTGCAGCTTCCGAGCGCTCACAGAGATCTTGAAGGAGGG 62
    |||||
QY 153 ValLeuIleGlnArgAsnProGlnLeuGluLeuLeuLeuLeuLeuLeuLeuLeu 172
    |||||
Db 63 GTCCTGATCCAGCGGAACCCAGCTCGCTACACAGACAGATTTTGTGAAGGACATC 122
    |||||
QY 173 PheHisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCys 192
    |||||
Db 123 TTCCACAGAGAGGACGCTGGCTCTACACTGATAGACACCAACCGCTCTCGGGCCTGC 182
    |||||
QY 193 HisProCysSerProMetCysLysGlySerArgCysTrpGlyGluSerSerGluAspCys 212
    |||||
Db 183 CACCCTCTTCTCCGATGTGAAGGGCTCCCGCTGCTGGGGAGAGGGTCTCTGAGGATTGT 242

```

```

213 GlnSerLeuThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuPro 232
    |||||
Db 243 CAGAGCCTGACGCGACACTGTCTGTCCCGTGGCTGTGCGCGTCAAGGGCCACTGGCC 302
    |||||
QY 233 ThrAspCysCysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCys 252
    |||||
Db 303 ACTGACTGCTGCCATGAGCAGTGTCTGCCGGTGCACGGGCCCAAGCACTCTGACTGC 362
    |||||
QY 253 LeuAlaCysLeuHisPheAsnHisSerGlyIleCysGluLeuHisCysProAlaLeuVal 272
    |||||
Db 363 CTGGCTGCCCTCCACTTCAACACAGTGGCATCTGTGAGTGCATCGCCAGCCCTGGTC 422
    |||||
QY 273 ThrTyrAsnThrAspThrPheGluSerMetProAsnProGluGlyArgTyrThrPheGly 292
    |||||
Db 423 ACCTACACACAGACACGCTTTGAGTCCATGCCAATCCGAGGGCGGTATACATTCCGC 482
    |||||
QY 293 AlaSerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySerCysThr 312
    |||||
Db 483 GCCAGCTGTGACTGCTGCTCCCTACAACTACCTTTCTAGCGAGTGGGATCCTGCACC 542
    |||||
QY 313 LeuValCysProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgCysGlu 332
    |||||
Db 543 CTGCTCTCCCCCTGCACACCAAGAGTGCACAGAGATGGAGAACACACGGGTGTGAG 602
    |||||
QY 333 LysCysSerLysProCysAlaArgValCysTyrGlyLeuGlyMetGluHisLeuArgGlu 352
    |||||
Db 603 AAGTGCACAGCCCTGTGCCCGAGTGTCTATGCTGGGCATGGAGCACTGGGAGAG 662
    |||||
QY 353 ValArgAlaValThrSerAlaAsnIleGln-GluPheAla-GlyCysLysLys---IleP 371
    |||||
Db 663 GTGAGGGCAGTTACGAGTGCCAAATATCCAGAGAGTTTGCCTGGCTGCAAAAAGATCTTT 722
    |||||
QY 371 heGlySerLeuAlaPheLeuPro---GluSerPheAspGlyAspPro---AlaSerAsnT 389
    |||||
Db 723 GGGGAGCCTGCCAATTTCTGCCCGAAAACTTTTGATGGGGAACCAACCTCCCAACA 782
    |||||
QY 389 hrAlaPro---LeuGlnProGlu 395
    |||||
Db 783 CTGGCCCCCGGTTCAGGCCCAAA 805

RESULT 11
BE746725
LOCUS      BE746725.1 GI:10160717
DEFINITION mRNA sequence.
VERSION     BE746725.1
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   NIH-MGC http://mgc.nci.nih.gov/.
            1 (bases 1 to 894)
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished
JOURNAL
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTF
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: L1CM758 row: e column: 16
            High quality sequence start: 3
            High quality sequence stop: 762.
            Location/Qualifiers
                1..894
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:3927927"

```

/tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_9"

/notes="Organ: ovary; Vector: pORB7; Site:1; XhoI; Site:2;
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 172 a 285 c 252 g 185 t
 ORIGIN

Alignment Scores:

Pred. No.: 5,11e-96 Length: 894
 Score: 1293.00 Matches: 261
 Percent Similarity: 85.76% Conservative: 4
 Best Local Similarity: 84.47% Mismatches: 29
 Query Match: 35.64% Indels: 17
 DB: 10 Gaps: 6

SEQ3 (1-653) x BE746725 (1-894)

Qy 230 ProLeuProThrAspCysCysHisGluGlnCysAlaAlaGlyCysThrGlyProLysHis 249
 Db 2 CCACTGCCACTGACTGCTGCCATGACAGTGTGCTGCCGCTGCACGGGCCCAAGCAC 61
 Qy 250 SerAspCysLeuAlaCysLeuHisPheAsnHisSerGlyIleCysGluLeuHisCysPro 269
 Db 62 TCTGACTGCCCTGCCCTCCACTTCAACCACAGTGGCATCTGTGAGTGCCTGCCCA 121
 Qy 270 AlaLeuValThrTyrAsnThrAspThrPheGluSerMetProAsnProGluGlyArgTyr 289
 Db 122 GCCCTGGTCACTACACACAGACAGCTTTGAGTCCATGCCCAATCCCGAGGCCGGTAT 181
 Qy 290 ThrPheGlyAlaSerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGly 309
 Db 182 ACATTGGCGCCAGCTGTGTGACTGTGCTGCTCCCTACAACTACCTTTCTACGAGCTGGA 241
 Qy 310 SerCysThrLeuValCysProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGln 329
 Db 242 TCTGCACTGCTGTGCTGCCCTTGCACCAACAGAGGTGACACGAGGATGGAACACAG 301
 Qy 330 ArgCysGluLysCysSerLysProCysAlaArgValCysTyrGlyLeuGlyMetGluHis 349
 Db 302 CGGTGTGAGAAGTGCACAGACCCCTGTGCCGAGTGTGTATGCTGGCATGGAGCAC 361
 Qy 350 LeuArgGluValArgAlaValThrSerAlaAsnIleGlnGluPheAlaGlyCysLysLys 369
 Db 362 TTGCGAGAGTGGAGGCGAGTTACAGTGCCTCAATATCCAGAGTTCCTGCTGCTCAAGAAG 421
 Qy 370 IlePheGlySerLeuAlaPheLeuProGluSerPheAspGlyAspProAlaSerAsnThr 389
 Db 422 ATCTTTGGGAGCTGGCATCTTCGCGAGAGCTTTGATGGGACCCAGCTCCACACT 481
 Qy 390 AlaProLeuGlnProGluGlnLeuGlnValPheGluThrLeuGluGluLeuThrGlyTyr 409
 Db 482 GCCCGCTCCAGCAGCAGCAGCTCCAAAGTGTGTTGAGACTCTGGAAGATCACAGTTAC 541
 Qy 410 LeuTyrIleSerAlaTrpProAspSerLeuProAspLeuSerValPheGlnAsnLeuGln 429
 Db 542 CTATACATCTCAGCATGGCGGACAGCTGCTGCTGAGCTGAGCTGCTTCCAGAACCTGCA 601
 Qy 430 ValIleArgGlyArgIleLeuHisAsnGlyAlaTyrSerLeuThrLeuGlnGlyLeuGlu 449
 Db 602 GTATCCGGGAGCAATCTGCACATGGCGCTACTCTGCTGACCTGCAAGTGTGGG 661
 Qy 449 YlleSer-TrpLeuGlyLeuArgSerLeuArg---GluLeu-GlySerGlyLeuAlaLeu 467
 Db 662 CATCAGCTGGCTGGGCTCGCTCACTTCACTGAGTGGAACTTGGGCGAGTGGCTGGCC--- 718
 Qy 468 IleHisAsnThrHisLeuCysPheValHisThrValProTrpAspGlnLeuPheArg 487

Db 719 TCATCATAATAAACCCACTC-TGTTGCTGCCCCACGCTGCCCTGGACCACTCTTCGG- 776
 Qy 488 AsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGluAspGluCysValGlyGlu 507
 Db 777 AACCCAGCAGCTCT---GTCCACATTGCCAACCCGCGAGGAGC----- 818
 Qy 508 GlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTyr-----GlyProGlyPro 525
 Db 819 ---ATTTTTCAGAGGCCCTGCTGCCCCAGTGTCCCGCTGCTTGTGGTCCGCCCC 875
 Qy 526 ThrGlnCysValAsnCysSer 532
 Db 876 ---CCGTGCTGGCGGCTCT 893

RESULT 12

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CA328613 795 bp mRNA linear EST 27-NOV-2002
 UI-M-FY0-cda-e-16-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
 IMAGE: 6826841 5', mRNA sequence.

CA328613

CA328613.1 GI:24546711

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 795)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabps@mail.nih.gov

Tissue procurement: Dr. Jim Lin, University of Iowa

cDNA library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..795

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6826841"

/tissue_type="whole brain"

/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_FY0"

/notes="Organ: Brain; Vector: pYX-Asc; Site:1: EcoR I;

Site:2: Not I; The library was constructed according

Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is AGCGACACAG. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

BASE COUNT 174 a 238 c 216 g 165 t

ORIGIN

Alignment Scores:

Pred. No.:	6.63e-89	Length:	795
Score:	1205.50	Matches:	219
Percent Similarity:	87.55%	Conservative:	13
Best Local Similarity:	82.64%	Mismatches:	31
Query Match:	33.23%	Indels:	2
DB:	14	Gaps:	1

SEQ3 (1-653) x CA328613 (1-795)

Qy	86	IleAlaHisAsn-GlnValArgGlnValProLeuGlnArgLeuArgIleValArgGlyTh	105
Db	3	ATCGCTCAACACGAGTGAACACAGCTCCCACTCGAGAGTTGCGATCGTCGAGAGGAC	62
Qy	105	rGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspProLeuAsnAs	125
Db	63	TCAGCTCTTTGAGGACAAGTATGCCCTGGCTGTGCTAGACAACCGAGACCTTTGGACAA	122
Qy	125	n---ThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeuArgSe	144
Db	123	CGTCACCAACCGCGCCCGCAGCAGACAACCCAGAAAGGCTGCGGGAGCTGCAGCTTCCGAAG	182
Qy	144	rLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeuCysTyrG	164
Db	183	TCTCACAGAGATCTTGAAGGAGGAGTTTGATCCGTGGGAACCCTCAGCTCTGCTACCA	242
Qy	164	nAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThrLeuI	184
Db	243	GGACATGGTTTTGTGGAGGATGTCTCCGTAGAATAACCACTGGCTCCTGTGCGACAT	302
Qy	184	eAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySerArgCy	204
Db	303	GGACACCAATCGTTCGCCGGCTGTCCACCTGTGTGCCCAACCTGCCAAAGACAATCACGTG	362
Qy	204	sTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAlaGlyClyCy	224
Db	363	TTGGGGTGAGAGTCTCTGAAGACTCTCAGACTTGACTGGCACCACCTGTAGTAGTGGCTG	422
Qy	224	sAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAlaAlaGlyCy	244
Db	423	TGCCGGGTGCAAGGGCCGGCTGCCACCTGACTGTTGCCATGAGCAGTGTGCTCAGGCTG	482
Qy	244	sThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGlyIleCy	264
Db	483	CACGGGTCCCAAGCATTCGTACTGCTCGGCTGCTCCACTTCAATCATAGTGGTATCTG	542
Qy	264	sGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSerMetProAs	284
Db	543	TCAGCTGCATGCCCGGCCCTTCACCTACACACACAGACACCTTCGAGTCCATGCTCA	602
Qy	284	nProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsnTyrLe	304
Db	603	CCCTCAGGGTCGTACACCTTTGTTGGTCAGCTGTGTGACCACCTGCCCCCTACAACTACCT	662
Qy	304	uSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnGluValThrAl	324
Db	663	CTTCCACGGAAGTGGATGCTCTGCATCTGGTCTGTCNCCCGAACAACCAAGAGGTCCACAG	722
Qy	324	aGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArgValCysTyrG	344
Db	723	TGAGACCGGAACACAGCGGTGTGAGAAATCGACAAGCCCTGTGCTGGAGTATGCTATGG	782
Qy	344	yLeuGlyMetGlu	348
Db	783	TCTGGCGATGGAG	795

RESULT	13
BX402419	
LOCUS	1201 bp mRNA linear EST 13-MAY-2003
DEFINITION	BX402419 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODIOJ032YB05 5-PRIME, mRNA sequence.
ACCESSION	BX402419
VERSION	BX402419.1 GI:30632074
KEYWORDS	EST.

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1201)
TITLE	Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL	Full-length cDNA libraries and normalization
COMMENT	Unpublished Contact: Genoscope

Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8568.f For more information about this cluster, see
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSIAI008ZH03QP1&cluster=8568.f)
[cgi-bin/cluster.cgi?seq=CSIAI008ZH03QP1&cluster=8568.f](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSIAI008ZH03QP1&cluster=8568.f). Contact :
Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSIAI008ZH03QP1.

```

FEATURES
    source
        Location/Qualifiers
            1..1201
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CSODI032YB05"
                /tissue_type="PLACENTA COT 25-NORMALIZED"
                /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                /note="1st strand cDNA was primed with a NotI-oligo(dT)
                    primer. Five prime end enriched, double-strand cDNA was
                    digested with Not I and cloned into the Not I and EcoR V
                    sites of the pcMVSPORT 6 vector. Library was normalized."
            235 a 376 c 321 g 223 t 46 others
BASE COUNT

```

ORIGIN		Alignment Scores:			
Pred. No.:	Score:	9.75e-88	Length:	1201	
Percent Similarity:	1194.50	Matches:	263		
Best Local Similarity:	52.18%	Conservative:	0		
Query Match:	52.18%	Mismatches:	11		
DB:	32.92%	Indels:	231		
	13	Gaps:	3		
SEQ3 (1-653) x BX402419 (1-1201)					
Qy	1	MetGluLeuAlaAlaLeuCysArgTrp-GlyLeuLeuLeuAlaLeuLeuProGlyAla	20		
Db	192	ATGGAGCTGGCGGCGCTTGGCGCTGGYGGKCTCCTCTCGCCCTCTGTGCCCGCCGAGC	251		
Qy	20	aAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProG1	40		
Db	252	CGGAGCACCCCAAGTGTCCACGGGCACAGACATCAAGCTGGGCTCCCTGCCAGTCCCGA	311		
Qy	40	uThrHisLeuAspMetLeuArgHisLeuTyrGlnGlnCysGlnValValGlnGlyAsnLe	60		
Db	312	GACCACCTGGACATGCTCCGCCACCTCTACCAAGGCTGCAGTGGTGCAGGAAACCT	371		
Qy	60	uGlnLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIuVa	80		
Db	372	GGAACCTCACCTACCTGCCGCCCAATGCCAGCTGTCTTCTTCCA-GAC-----	419		
Qy	80	lGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuAr	100		
Db	419	-----	419		
Qy	100	gIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGl	120		
Db	419	-----	419		
Qy	120	yAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyLeuArgGluLe	140		
Db	419	-----	419		

```

Qy 140 uGlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProG1 160
Db 419 -----
Qy 160 nLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAl 180
Db 419 -----
Qy 180 aLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLy 200
Db 419 -----
Qy 200 sGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCy 220
Db 419 -----
Qy 220 sAlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCy 240
Db 419 -----
Qy 240 sAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHi 260
Db 419 -----
Qy 260 sSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheG1 280
Db 419 -----
Qy 280 uSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPr 300
Db 419 -----
Qy 300 oTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnG1 320
Db 420 ----AACTACTCTTCTACGGACGTGGGATCTCGTCCACCTCGTGGCCCTGGCACACCA 475
Qy 320 nGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaAr 340
Db 476 AGAGGTGACACGACGAGGATGGACACACACGCGTGTGAGAAGTGCAGCAAGCCCTGTGCCG 535
Qy 340 qValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAs 360
Db 536 AGTGCTGTAGGCTGGGATGAGACACTTGGCAGAGGTGAGGACAGTACCAGTGCCAA 595
Qy 360 nIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSe 380
Db 596 TATCCAGGAGTTCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAG 655
Qy 380 rPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPh 400
Db 656 CTTTGTATGGGACCCAGCCTCCAACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 715
Qy 400 eGluThrLeuGluGluThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPr 420
Db 716 TGAGACTCTGGAAGAGATCACAGGTACCTATACATCTACATGCGCGGACAGCCCTGCC 775
Qy 420 oAspLeuSerValPheGlnAsnLeuGlnValIleArg-GlyArgIleLeuHisAsnGlyA 440
Db 776 TGACCTCAGGCTCTCCAGAACCTGCAAGTAATCCGGGGACCAATTCGCACAATGGCG 835
Qy 440 laTyrSerLeuThrLeuGln-GlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArg 459
Db 836 CCTACTCGCTGACCTGCAAGGCTGGGGCATCAGCTGGCTGGGGCTCGCTCACTGAGG 895
Qy 460 GluLeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThr 479
Db 896 GAATGGGCAAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 955
Qy 480 ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg 499
Db 956 GTGCCCTGGGACARCTCTTTC---GGNACCGCACAGCTCTGTC---ACATGCCACCGG 1009
Qy 500 ProGlu 501

```

```

Db 1010 CCAGAG 1015
RESULT 14
AW370693
LOCUS
DEFINITION
QV1-BT0260-011199-024-all BT0260 Homo sapiens cDNA, mRNA sequence.
ACCESSION
AW370693
VERSION
AW370693.1 GI:6875347
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 614)
HCGP http://www.ludwig.org.br/ORESTES.
AUTHORS
The FAPESP/LICR Human Cancer Genome Project
TITLE
Unpublished
JOURNAL
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV1&t2=QV1-BT0260-
011199-024-all&t3=1999-11-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 612.
FEATURES
source
Location/Qualifiers
1..614
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BT0260"
/note="Organ: breast; Vector: puc18; Site1: SmaI; Site2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 119 a 198 c 172 g 124 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1,09e-83 Length: 614
Score: 1140.00 Matches: 197
Percent Similarity: 98.99% Conservative: 0
Best Local Similarity: 98.99% Mismatches: 2
Query Match: 31.42% Indels: 0
DB: 9 Gaps: 0
SEQ3 (1-653) x AW370693 (1-614)
Qy 153 ValLeuIleGlnArgAsnProGlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIle 172
Db 17 GTCTTGATCCAGCGGAACCCAGCTCTGCTACAGACACAGATTTGTGGAAGGACATC 76
Qy 173 PheHisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCys 192
Db 77 TTCACAAGAACACACAGCTGGCTCTCACACTGATAGACACCAACCGCTCTCGGGCCTGC 136
Qy 193 HisProCysSerProMetCysLysGlySerArgCysTrpGlyGluSerSerGluAspCys 212
Db 137 CACCCCTGTCTCCGATGTGAAGGCTCCCGCTCTGGGGAGAGAGTTCTGAGGATGT 196
Qy 213 GlnSerLeuThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuPro 232

```

```

197 CAGAGCCTGACGGCAGCTCTCTGCGCGTGGCTGTGCGCGTCAAGGGCGCACTGCGCC 256
233 ThrAspCysCysHisGluGlnCysAlaAlaGlyCysThrGlyProIyHisSerAspCys 252
257 ACTGACTGCTGCATGAGCAGTGTGTCGGGTGTCACGGGCCCAAGCAGCTGACTGC 316
253 LeuAlaCysLeuHisPheAsnHisSerGlyIleCysGluLeuHisCysProAlaLeuVal 272
317 CTGGCCTGCTCCACTTCAACACACAGTGGCATCTGTGAGCTGCAGTCCGACCCCTGGTC 376
273 ThrTyrAsnThrAspThrPheGluSerMetProAsnProGluGlyArgTyrThrPheGly 292
377 ACCTACACACAGACACGTTTGTGTCATGCCAATCCCGAGGGCGGTATACATTCGGC 436
293 AlaSerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySerCysThr 312
437 GCAGCTGTGACTGCTCCCTCCCTACAACTACCTTCTACGAGCGTGGATCCCTGCACC 496
313 LeuValCysProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgCysGlu 332
497 CTGCTGCTGCCCCGTCACAAACAGAGGTGACAGCAGAGGATGGAACACAGCGGTGGG 556
333 LysCysSerIyProCysAlaArgValCysTyrGlyLeuGlyMetGluHisLeuArg 351
557 AAGTGCAGCAAGCCCTGTTGCGGAGTGTGTATGTTCTGGGATGGAGCACTTGCGA 613

```

```

RESULT 15
AW410534
LOCUS
DEFINITION
  AW410534
  mRNA sequence.
  613 bp mRNA linear EST 29-JUN-2000
  fh06h06.x1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2961635 5',
  mRNA sequence.
  AW410534
  EST.
  AW410534.1 GI:6936075
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  NIH-MGC http://mgi.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-remail.nih.gov
  Tissue procurement: ATCC
  CDNA Library Preparation: Ling Hong/Rubin Laboratory
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: National Institutes of Health Intramural
  Sequencing Center (NISC)
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Plate: LLC56 row: O column: 12
  Seq primer: -21M13 forward primer (ABI).
  Location/Qualifiers
    1. .613
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:2961635"
      /tissue_type="rhabdomyosarcoma"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH_MGC_17"
      /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
      Site_2: XhoI; cDNA made by oligo-dr priming.
      Directionally cloned into EcoRI/XhoI sites using the
      following 5' adaptor: GGCAGGAG(G). Size-selected >500bp
      for average insert size 1.8kb. Library constructed by
      Ling Hong in the laboratory of Gerald M. Rubin (University
      of California, Berkeley) using ZAP-cDNA synthesis kit
      (Stratagene) and Superscript II RT (Life Technologies)."
    109 a 205 c 185 g 113 t 1 others

```

FEATURES

source

```

ORIGIN
Alignment Scores:
Pred. No.: 1,91e-83 Length: 613
Score: 1137.00 Matches: 195
Percent Similarity: 97.99% Conservative: 0
Best Local Similarity: 97.99% Mismatches: 4
Query Match: 31.34% Indels: 0
DB: 9 Gaps: 0

SQ03 (1-653) x AW410534 (1-613)
QY 455 LeuArgSerLeuArgGluLeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeu 474
Db 5 CTGCGCTCCTCCTGAGGGAAGTGGGAGTGGAGTGGCCCTCATCCACCAATACACCCACCTC 64
QY 475 CysPheValHisThrValProTyrAspGlnLeuPheArgAsnProHisGlnAlaLeuLeu 494
Db 65 TGCTTTCGTGCACACGCGGCCCTGGGACCACTCTTTTCGGAACCCGACCAAGCTCTGCTC 124
QY 495 HisThrAlaAsnArgProGluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeu 514
Db 125 CACACTGCCAACCGGCCAGAGAGTGTGTGGCGAGGGCTGGCTGCCACCAAGCTG 184
QY 515 CysAlaArgGlyHisCysTyrProGlyProGlyProThrGlnCysValAsnCysSerGlnPhe 534
Db 185 TGGCGCCCGAGGGCACTGCTGGGGTCCAGGGCCCACTAGTGTCACTGACAGCCAGTTTC 244
QY 535 LeuArgGlyGlnGluCysValGluCysArgValLeuGlnGlyLeuProArgGluTyr 554
Db 245 CTTCGGGCGCAGGAGTGGGTGAGGAATGCCAGTACTCGAGGGCTGCCACAGGAGTAT 304
QY 555 ValAsnAlaArgHisCysLeuProCysHisProGluCysGlnProGlnAsnGlySerVal 574
Db 305 GTGAATGCCAGCACTGTTTGGCGTGGCCACCTGAGTGTACGCCCCCAAGTGGCTCAGTG 364
QY 575 ThrCysPheGlyProGluAlaAspGlnCysValAlaCysAlaHisTyrIyLysAspProPro 594
Db 365 ACCCTGTTTGGACCGGAGGCTGACCAGTGTGTGGCGCTGTGCCCACTATAAGGACCCCTCCC 424
QY 595 PheCysValAlaAlaArgCysProSerGlyValIyLysProAspLeuSerTyrMetProIleTrp 614
Db 425 TTCTGCGTGGCGCGCTGCCCGCAGCGGTGTGAACCTGNACCTCTCTACATGCCCATCTGG 484
QY 615 LysPheProAspGluGluGlyAlaCysGlnProCysProIleAsnCysThrHisSerCys 634
Db 485 AAGTTTCCAGATGAGAGGGCGCATGCCAGGCTTGCCCCCATCACTGACCCACCTCTCTGT 544
QY 635 ValAspLeuAspAspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 545 GTGGACCTGGATGACAAGGGTGTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTC 601

```

Search completed: October 16, 2003, 17:04:35
Job time : 5138.68 secs

BASE COUNT

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 11:08:07 ; Search time 675.009 seconds
(without alignments)
2540.503 Million cell updates/sec

Title: SEQ3

Perfect score: 3628

Sequence: 1 MELAALCRWGLLLALLPPGA.....CVDLDRKCPAEQRASPLTS 653

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE1=frame+p2n_model -DEV=xlp
-O=/cgn2_1/USP70_spool_p/HOLLERAN480/runat_15102003_131915_20562/app_query.fasta_1.4685
-DB=PublishedApplications_NA -QFM=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NAXLEN=2000000000 -USER=HOLLERAN480 -CGN_1_1_397 @runat_15102003_131915_20562
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3628	100.0	3765	12	US-10-207-498-5 Sequence 5, Appli

2	3628	100.0	3768	9	US-09-811-123-8	Sequence 8, Appli
3	3628	100.0	3768	9	US-09-811-115-2	Sequence 2, Appli
4	3628	100.0	3768	10	US-09-854-356-9	Sequence 9, Appli
5	3628	100.0	3768	10	US-09-930-125-1	Sequence 1, Appli
6	3628	100.0	3768	12	US-10-313-644-1	Sequence 1, Appli
7	3628	100.0	4473	11	US-09-441-411-5	Sequence 5, Appli
8	3628	100.0	4473	12	US-10-101-510-81	Sequence 81, Appli
9	3628	100.0	4473	14	US-10-146-473-32	Sequence 32, Appli
10	3628	100.0	4473	14	US-10-207-655-44	Sequence 44, Appli
11	3628	100.0	4530	10	US-09-877-177-11	Sequence 11, Appli
12	3628	100.0	4530	12	US-10-007-926A-119	Sequence 119, App
13	3628	100.0	4530	12	US-10-101-510-124	Sequence 124, App
14	3628	100.0	4530	12	US-10-338-730-1	Sequence 1, Appli
15	3628	100.0	4530	14	US-10-177-293-125	Sequence 125, App
16	3628	100.0	4606	12	US-09-971-392-70	Sequence 70, Appli
17	3628	100.0	4642	14	US-10-198-846-10896	Sequence 10896, A
18	3628	100.0	9274	9	US-09-811-123-7	Sequence 7, Appli
19	3628	100.0	9274	9	US-09-811-115-1	Sequence 1, Appli
20	3585	98.8	4543	10	US-09-769-508-1	Sequence 1, Appli
21	3106	85.6	3955	10	US-09-870-759-117	Sequence 117, App
22	3106	85.6	3955	10	US-09-854-356-10	Sequence 10, Appli
23	3106	85.6	3955	12	US-09-751-708A-117	Sequence 117, App
24	3100.5	85.5	3771	10	US-09-854-356-11	Sequence 11, Appli
25	1644.5	45.3	2091	9	US-09-821-883-9	Sequence 9, Appli
26	1630.5	44.9	2070	9	US-09-821-883-7	Sequence 7, Appli
27	1614.5	44.5	1692	9	US-09-821-883-8	Sequence 8, Appli
28	1612	44.4	1665	9	US-09-821-883-6	Sequence 6, Appli
29	1608	44.3	1437	9	US-09-821-883-10	Sequence 10, Appli
30	1587	43.7	867	9	US-09-821-883-24	Sequence 24, Appli
31	1532.5	42.2	3633	9	US-09-725-433-1	Sequence 1, Appli
32	1532.5	42.2	5532	12	US-10-007-926A-137	Sequence 137, App
33	1532.5	42.2	5532	12	US-10-101-510-95	Sequence 95, Appli
34	1530.5	42.2	2643	14	US-10-172-620-15	Sequence 15, Appli
35	1530.5	42.2	10058	10	US-09-974-298-98	Sequence 98, Appli
36	1518.5	41.9	5264	10	US-09-920-300A-1731	Sequence 1731, Ap
37	1518.5	41.9	5264	12	US-10-099-926-1731	Sequence 1731, Ap
38	1518.5	41.9	5264	13	US-10-033-528-1731	Sequence 1731, Ap
39	1451	40.0	5484	10	US-09-940-101-1	Sequence 1, Appli
40	1451	40.0	5484	14	US-10-207-655-46	Sequence 46, Appli
41	1447	39.9	2601	10	US-09-940-101-3	Sequence 3, Appli
42	1435	39.6	4026	12	US-10-207-498-1	Sequence 1, Appli
43	1435	39.6	4879	12	US-10-101-510-137	Sequence 137, App
44	1435	39.6	4879	12	US-10-388-410-6	Sequence 6, Appli
45	1435	39.6	4879	14	US-10-172-620-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1

US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREULIN AND HER3
; FILE REFERENCE: 30448.103-US-01
; CURRENT APPLICATION NUMBER: US/10/207.498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)

US-10-207-498-5

Alignment Scores:

Pred. No.: 0 Length: 3765
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ3 (1-653) x US-10-207-498-5 (1-3765)

Qy	1	MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGluYala	20
Db	1	ATGAGAGTGGCGGCTTGTGCGCTGGGGCTCCTCGCCCTTGGCCCGGAGCC	60
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	61	CGGAGCACCCAAAGTGTGACCGGCACAGATGAAGCTGGGGCTCCTGCGAGTCCCGAG	120
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
Db	121	ACCCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGTGGTGCAGGGAACTG	180
Qy	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	181	GAACCTACCTACCTGCGCCACCAATGCCAGCTGTCTTCTGTCAGGATATCCAGGAGTG	240
Qy	81	GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	241	CAGGGCTACGTGCTTCATCGCTTCACCAACAGTAGGAGGCTGCCACCTGCAGAGGCTCGG	300
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	301	ATTGTCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGTAGACAAATGA	360
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	361	GACCGCTGAAATACACCCCTGTCTCACAGGGGCTCCCGAGGAGGCTGCGGGAGCTG	420
Qy	141	GlnLeuArgSerLeuThrGluLeuLeuLysGlyGlyValLeuLeuGlnArgAsnProGln	160
Db	421	CAGCTTGAAGCCCTCACAGATCTTGAAGGAGGGCTCTTGATCCAGCGGAACCCCGAG	480
Qy	161	LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	481	CTCTGCTACCAAGGACAGATTTTGTGAAGGACATCTTCCACAAAGAAACACCACTGGCT	540
Qy	181	LeuThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	541	CTCACACTGATAGACACCAACCGCTCTCGGGGCTGCCACCCCTGTCTCCGATGTGAAG	600
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	601	GGCTCCCGCTGCTGGGAGAGAGTCTTGAGGATTCAGAGCTTCAGAGCTGACCGCAGCTGT	660
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	661	GCCGGTGGCTGTGCCCTGCAAGGGGCCACTGCCACTGACTGCTGCCATGACAGCTGT	720
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	721	GCTCCCGGCTGCAGGGGCCCAAGCACTCTGACTGCTGGCTGCCCTGCCCTTCAACAC	780
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATGTGAGCTGCACTGCCAGCCCTGCTGCTACCTACACACAGACAGCTTTGAG	840
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCAATCCGAGGGCCGGTATACATTCGGGCCAGCTGTGTGACTGCTGTGCC	900
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320

Db	901	TACAACTACCTTTCTACGGAGCTGGATCTGTCACCTCTGCTGCTGCCCTGCACAAACCA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1020
Qy	341	ValCystyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGGTCTGGGCATGGACACTTGGCAGAGGTGAGGAGCTTACAGTGCCCAAT	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGGAGTTGTGTGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGATGGGACCCAGCCTCCACACTGCCCGCTCCAGCAGAGAGCTCCCAAGTGTGT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCAGAGTTACTATACATCTCAGCATGCCCGGACAGCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGCTCTTCCAGAACCTGCAAGTAATCCGGGAGCAATCTGCACATGGCGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGCTCCTCACTGAGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGCAGTGGACTGGCGCTCATCCACCAATAACACCCACCTCTGCTTCGTCACACGCTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGTCTCCACACTGCCAACCGCCA	1500
Qy	501	GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGTGTGTGGCGAGGGCTGGCTGCCCTGCCAGCTGCGCGCCGAGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGTGCCAGGGCCCAACCCAGTGTCAACTGCAGCCAGTTCTCTCGGGGCCAGGAGTGC	1620
Qy	541	ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTGAGTGTACGCCCCAGAAAGGCTCAGTACCTGTTTGGACCGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCAGTGTGTGGCTGTGCCCACTATAAGGACCCCTCCCTTCTGCTGCGCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTGTTCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Db	1861	GGCGCATGCCAGCCTTGCCCATCACTGCACCCACCTCTCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGAGGAGGAGCCAGCCCTCTGACGTCC	1959

RESULT 2

US-09-811-123-8

Sequence 8, Application US/09811123
 Patent No. US20020001587A1

GENERAL INFORMATION:

APPLICANT: Sharon Erickson
 APPLICANT: Ralph Schwalli
 APPLICANT: Mark Sliwkowski
 TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
 TITLE OF INVENTION: ANTIBODY-MAYTANSINOID CONJUGATES
 FILE REFERENCE: GENE.073A2
 CURRENT APPLICATION NUMBER: US/09/811,123
 CURRENT FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/238,327
 PRIOR FILING DATE: 2000-10-05
 PRIOR APPLICATION NUMBER: 09/602,530
 PRIOR FILING DATE: 2000-06-23

NUMBER OF SEQ ID NOS: 11

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8

LENGTH: 3768

TYPE: DNA

ORGANISM: Homo sapiens

US-09-811-123-8

Alignment Scores:

Pred. No.: 0 Length: 3768
 Score: 3628.00 Matches: 653
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

SEQ3 (1-653) x US-09-811-123-8 (1-3768)

QY 1 MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 DB 1 ATGGAGCTGGCGCTTGTGGCGCTGGGGCTCCTCCTCGCCCTCTTGGCCCCCGAGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 61 GCGAGCACCACCAAGTGTGACCGGCGACAGACATGAAGCTCGGGCTCCTCGCCAGTCCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60
 DB 121 ACCCAGCTGGACATGCTCGCCACCTCTACAGAGGCTGCCAGTGGTGCAGGGAACCTG 180
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 DB 181 GAACTCACCTACCTGCCCCACCAATGCCAGCCTGTCTTCTCGAGGATATCCAGGAGTG 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 241 CAGGGCTACGTGCTATCGCTCACAAACCAAGTGAGGCAAGTCCCACTGCAGAGGTGCGG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 301 ATTGTGCGAGGACCCAGCTCTTTCAGGACAACTATGCTCGCCGTGCTAGACAAATGGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB 361 GACCGGCTGAACAATACCAACCCCTGTACAGGGGCTCCCAAGAGGCTCGCGGAGTG 420
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValValLeuIleGlnArgAsnProGln 160
 DB 421 CAGCTTCGAAGCCTCACAGAGATCTGAAGAGAGGGTCTTGATCCAGCGGAACCCACG 480
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 DB 481 CTCTGCTACAGACACGATTTTGTGGAGGACATCTTCCACAAAGAACACACCGCTGGCT 540
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600

QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 601 GGCTCCCGCTGCTGGGAGAGAGATTCTGAGGATTGTACAGAGCTGACGGCACTGTCTGT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
 DB 661 GCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGTGCTGCTGCTGCTGCTG 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 721 GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 781 AGTGGCATCTGTGAGCTGCACGTGCCAGCCCTGGTACCTCAACACACACACAGCTTGAG 840
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 841 TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGGGCCAGCTGTGTGACTGCTGCTGCT 900
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 901 TACAACCTACCTTCTACGGAGCTGGGATCCTGCACCTCTGCTGCCCTGCACACCAACAA 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCCGA 1020
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1021 GTGTCTATGTCTGGCATGGAGCACCTTGCAGAGGTGGGGCAGTTACCAAGTGCCTT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1081 ATCCAGGAGTTGTGCTGCTGCAAGAAGATCTTTGGAGCTGGCATTTCTGCGGAGAGC 1140
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB 1141 TTTGATGGGACCCAGCTTCCAAACCTGCCCGCTCCAGCCAGAGCAGCTCCCAAGTGT 1200
 QY 401 GluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
 DB 1201 GAGACTTGAAGACAGATCACAGTTACCTATACATCTCAGCATGCGCGGACAGCTGCT 1260
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1261 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGACGAATCTTGCAATGCGGCC 1320
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTyrLeuGlyLeuArgSerLeuArgGlu 460
 DB 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCATGCTGGCTGGGCTGCTGCTGCTGCTGCT 1380
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1381 CTGGCAGTGAGCTGGCCCTCATCCACCATACACCCACTCTGCTGCTGCTGCTGCTGCTG 1440
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuLeuHisThrAlaAsnArgPro 500
 DB 1441 CCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCA 1500
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1501 GAGGACGAGTGTGGGAGGGCTGGCTGCCACCAAGCTGTGCCCGGAGGCGACCTGTC 1560
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1561 TGGGTCCAGGGCCCAAGCTGTCACTGACCCAGTTCCTTCTGCGGGCCAGGAGTGC 1620
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGlyTyrValAsnAlaArgHisCys 560
 DB 1621 GTGGAGGAATGCCAGTACTGCGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580

```

|||||
Db 1691 TTGGCGTCCCACTGAGTGTACGCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1740
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAraCys 600
|||||
Db 1741 GCTGACCAGTGTGGCGCTGTGCCCATATPAAGGACCTCCCTTCTCGGTGGCCCGGTGC 1800
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProLleTyrLysPheProAspGluGlu 620
|||||
Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
|||||
Db 1861 GGCGCATGCCAGCCTTGCCCATCAACTGCACCCACCTCTGTGTGGACCTGGATGACAAG 1920
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
|||||
Db 1921 GGCTGCCCGCGGACGAGAGACCGCCTCTGAGTCC 1959

RESULT 3
US-09-811-115-2
; Sequence 2, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred. No.: 0 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Db: 9 Gaps: 0

SEQ3 (1-653) x US-09-811-115-2 (1-3768)
Qy 1 MetGluLeuAlaAlaLeuCysArgTyrGlyLeuLeuAlaLeuLeuProGlyVala 20
|||||
Db 1 ATGAGCTGGCGGCTGTGGCGTGGGGCTCTCTCGCCCTTTGCCCGGAGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
|||||
Db 61 GCGAGCACCAAGTGTGCACCGGCACACATGAAGCTGGCGCTCCCTGCCAGTCCCGAG 120
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
|||||
Db 121 ACCACCTGGACATGCTCCCGCACCTTACAGGGCTGCCAGGTGGTGCAGGGAACCTG 180
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
|||||
Db 181 GAACCTACCTACCTGCCCAATGCCAGCTGTCTCTCTGAGGATATCCAGGAGGTG 240
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
|||||
Db 241 CAGGGCTACGTGCTCATCGCTCACAACCAAGTAGGAGGTCCTCCACTGCAGAGCTCGG 300
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
|||||

```

```

Db 301 ATTTGCGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCGGTCTAGACAATGGA 360
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
|||||
Db 361 GACCCGCTGAACAATACACCCCTGTACAGGGGCTTCCACAGAGAGGCTTGGGGAGCTG 420
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
|||||
Db 421 CAGCTTCAAGCCCTCAGAGATCTTGAAGAGGGGTCTTGATCCAGCGAACCCTGCT 480
Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
|||||
Db 481 CTCTGCTACCCAGGACAGATTTTGTGAAGGACATCTTCACAAGAACAACACCTGGCT 540
Qy 181 LeuThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
|||||
Db 541 CTCACACTGATAGACCAACCCGCTCTCGGGGCTGCCACCCCTGTCTCCGATGTGTAA 600
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
|||||
Db 601 GGCTCCCGCTGTGGGAGAGAGTTCTGAGGATTGTCAGAGCTGACGCGCACTGTCTGT 660
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCys 240
|||||
Db 661 GCCGGTGGCTGTGCCCTGCCAAGGGCCACTGCCCTGACTGCTGCCATGACAGCTGT 720
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
|||||
Db 721 GCTGCCGCTGCAGCGGCCCAACACTCTGACTGCTGGCTGCCCTCCACTTCAACAC 780
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
|||||
Db 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGCTACCTACAAACACAGACACCTTTGAG 840
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
|||||
Db 841 TCCATGCCCAATCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGCTGCTGCC 900
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAspGln 320
|||||
Db 901 TACAACTACCTTTCTACGGAGCTGGGATCTGCACCTCTGCTGCCCTCCCTGCACACCAA 960
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
|||||
Db 961 GAGGTGACAGCAGAGGATGGAACACACGCGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1020
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValAlaValThrSerAlaAsn 360
|||||
Db 1021 GTGTGCTATGCTGTGGCATGGAGCACTTGCAGAGGTGAGGGCAGTTACCAAGTSCCAAT 1080
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
|||||
Db 1081 ATCCAGAGTGTGCTGCTGCCAAGAAATCTTTGGGAGCCTGGCATTTCTGCCCGAGAGC 1140
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
|||||
Db 1141 TTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCAGAGAGTCCCAAGTGTGT 1200
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
|||||
Db 1201 GAGACTTGAAGAGATCAGAGGTACCTATACATCTCAGCATGGCGCGGACAGCTGCCT 1260
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyVala 440
|||||
Db 1261 GACCTCAGCTCTTCCAGAACCTTCAAGTAATCCGGGGAGCAATCTTCGACAAATGGCGCC 1320
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
|||||
Db 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGTGCGCTCACTCAGGGAA 1380
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
|||||
Db 1381 CTGGGCAGTGGACTGGCCCTCATCCACCAACACCACCTCTGCTTTCGTGTCACACGGTG 1440

```


Db	841	TCCATGCCCAATCCCGAGGCGCGGTATACATTGGCGCCAGCTGTGTGACTGCTGTGCC	900
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACTACCTTCTACGGACGTGGGATCTGCACCCCTCGTCTGCCCTGTGCACACCAA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACAGCAGAGGATGGAACACACACGCGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGGTCTGGGCATGTGGACACTTCGAGAGGTGAGGCGAGTTACCACTGTCCAAT	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGAGATTGCTGGCTCGAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGATGGGACCCAGCCTCAACACTGCCCGCTCCAGCGAGCAGCTCCAAAGTGT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCACAGGTTACTATACATCTCAGCATGGCGGACAGCGCTGCC	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGTCTCCAGAACCTCGCAAGTAAATCCGGGAGCAATCTGCACAATGGCGC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGGCTGCCTCCTCATGTAGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGGAGTGGACTGGCCCTCATCCACCAATAACACCCACCCTCTGCTGTGTGCACAGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACAGAGCTCTTTCGGAACCCGCACAGCTCTGTCTCCACACTGTCACACCGGCCA	1500
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACGAGTGTGTGGCGAGAGGCTGGCCTGCCACCAGCTGTGTGGCCCGAGGGCAGTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGGCCACCCAGTGTGTCAACTGCAGGCAGTCTCTTCTGGGGCCAGGAGTGC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGAAATGCCAGTACTGCAAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTGAGTGTGACGCCCAAGTGGCTCAGTGACCTGTGTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys	600
Db	1741	GCTGACCAGTGTGTGGCGCTGTGCCACTATAAGGACCTCTCCCTCTCGGTGGGCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCAGCGGTGTGAACCTGACCTCTCTACATCCCATCTGGAAGTTCCTCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCCCTTGGCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653

```
Db 481 CTCTGCTACCGACACGATTGTGGAAGGACATCTTCCACAAGACAACACGAGTGCT 540
QY 181 LeuThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTCTCCGATGTGAAG 600
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTGCTGGGGAGAGATTCTGAGGATTGTACAGACCTTGACCGGCACATGCTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCCGGTGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGTGTCATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTGCGGCTGACCGGGCCCAAGCACACTCTGACTGCCGTGGCTGCCTCCACTTCAACAC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGCATCTGTGAGCTGCACTGCCAGCCCTGGTGCCTACACACAGACACAGTTCGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCCGAGGGCGGTATATACATTTCGGCGCAGCTGTGTGACTGCCTGCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACACTCTTCTACCGAGCTGGGATCCCTGCACCCCTCGTCTGCCCTGCACAAACCA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGCAGAGATGGAAACACAGCGGTGTAGAGTGCAGCAACCCCTGTGCCCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1021 GTGTGCTATGCTGGGATGAGCAGCTTGGCAGAGGTGAGGGCAGTTACCAGTGCCCAAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 ATCCAGGAGTTGTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1141 TTTGATGGGACCCAGCCCTCCAACTGCCCGCTCCAGCCAGACGACTCCAAGTGT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
Db 1201 GAGACTCTGGAAGAGATCACAGGTATACCTATACATCTCAGCATGGCGGCAGACGCTGCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTCAGCGCTCTCCAGAAGCTGCAAGTAAATCGGGGACCAATTCGCAAAATGGCGCC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGGAA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGCAGTGGAGTGGCCCTCATCCACCATTAACACCCACCTCTGCTGTCACACAGGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCCTGGGACCAAGCTCTTTTCGAACCCGACCAAGCTCTGCTCCACACTGCCCAACCGGCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGACAGTGTGTGGGCGAGGCTGGCCCTGCCACACAGCTGTGGCCCGGAGGACACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
```

```
Db 1561 TGGGTCCAGGGCCCAACCCAGTGTGTCAACTGCAGCCAGTTCCTTTCGGGCCAGGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGAGGAATGCCAGTACTGCAGGGCTTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCGTGCACCCCTGAGTGTGAGCCCGCAGAAATGCTCAGTGACCTGTTTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1741 GCTGACCAAGTGTGGCTGTGCCCACTATAGGACCTTCCCTTCTGCTGGCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 1861 GCGCATGCCAGCTTGGCCCATCAACTGCACCCACTCCTGTGTGACCTGGATGACAAG 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959
```

RESULT 6

```
US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313,644
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-313-644-1
```

Alignment Scores:

Pred. No.:	0	Length:	3768
Score:	3628.00	Matches:	653
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

SEQ3 (1-653) x US-10-313-644-1 (1-3768)

```
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGAGCTGGCGGCGCTTGTGCGGTGGGGGCTCTCTCTCGCCCTTTCCTCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCACAAGTGTGCACCGCACACATGAAGCTGCGGCTCCCTGCGAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACCAGGCTGCCAGGTGGTGCAGGAAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerPheLeuLeuGlnAspIleGlnGluVal 80
```



```

Db      |||||GAACTCACTACCTGCGCCACCAATGCCAGCGTGCCTTCTTCGAGGATATCCAGAGGTG 240
Qy      |||||GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      |||||241 CAGGCTACGTGCTCTGCTCACACCAAGTAGGAGGCTGCCACTGCAGAGCTCGG 300
Qy      |||||101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      |||||301 ATTTGGCAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTCTAGACAAATGGA 360
Qy      |||||121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      |||||361 GACCCGCTGAACAATACACCCCTGTCACAGGGCCCTCCCCAGGAGCCTGCGGGAGCTG 420
Qy      |||||141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db      |||||421 CAGCTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG 480
Qy      |||||161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      |||||481 CTCGTCTACCAAGGACAGATTTCTGGAAGGACATCTTCACAAAGAACCAACAGCTGGCT 540
Qy      |||||181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      |||||541 CTCACACTGATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGAAG 600
Qy      |||||201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      |||||601 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATGTTCAGAGCCTGACGGCAGCTGTCTGT 660
Qy      |||||221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db      |||||661 GCCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720
Qy      |||||241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      |||||721 GCTGCCGGCTGCACGGGCCCAACGACTCTGACTGCTGGCCTGGCCTCCACTTCAACAC 780
Qy      |||||261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      |||||781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTGCACCTACAAACACAGACAGCTTTGAG 840
Qy      |||||281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      |||||841 TCCATGCCCAATCCCGAGGCCGGTATACATTCCGGCCAGCTGTGTGACTGCTGCTGCC 900
Qy      |||||301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      |||||901 TACAACCTACCTTCTACGGAGCTGGGATCTCGACCCCTGCTGTGCCCTCCCTGCACACCAA 960
Qy      |||||321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      |||||961 GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1020
Qy      |||||341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      |||||1021 GTGTGCTATGCTTGGCATGGAGCACTTCGAGAGGTGAGGCGAGTTACACAGTGCAT 1080
Qy      |||||361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      |||||1081 ATCCAGGAGTTGCTGGCTCAAGAGATCTTTGGGAGCTGGCATTTCTGCGCGAGAGC 1140
Qy      |||||381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      |||||1141 TTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGAGAGCTCCAAGTGT 1200
Qy      |||||401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db      |||||1201 GAGACTCGGAAGAGATCAAGGTTACCTATACATCTCAGCATGGCGGACAGCCGTGCCT 1260
Qy      |||||421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440

```

```

Db      1261 GACCTCAGCGCTTCTCCAGAAACCTCAAGTAATACGGGGACGAATCTTGCACAATGGCGCC 1320
Qy      441 TyrSerLeuThrLeuGlnGlyLeuGlyTleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1321 TACTCGCTGACCTTGCAGGGCTGGCATACAGCTGGCTGGGCTTGGGCTCTACTGAGGAA 1380
Qy      461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1381 CTGGGAGTGAGCAGTGGCCCTCATCCACATAAACACCACCTCTCTCTCGTCCACACGGTG 1440
Qy      481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db      1441 CCCTGGGACAGCTCTTTTCGGAACCCGACCACTCTGTCCACACTGCCAACCCGCCA 1500
Qy      501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db      1501 GAGGACAGCTGTGTGGGAGGGCTGGCCCTGCCACAGCTGTGCCGCCGAGGACACTGC 1560
Qy      521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db      1561 TGGGTCCAGGGCCCACTGAGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC 1620
Qy      541 ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db      1621 GTGGAGGAATGCCAGTACTGCAGGGCTCCCCAGGAGTATGTGAATGCCAGCACTGT 1680
Qy      561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db      1681 TTGCCGTGCCACCTGAGTGTACAGCCCAAGTGGCTCAGTACCTGTTTGGACCCGAG 1740
Qy      581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db      1741 GCTGACCAGTGTGGCTGTGCCACTATAAGAGACCTCCCTTCTCGTGGCCCGCTGC 1800
Qy      601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db      1801 CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
Qy      621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db      1861 GGGCATGCCAGCCTTGCCCACTCACTGCACCCACTCTCTGTGTGGACCTGGATGACA 1920
Qy      641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db      1921 GGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTCC 1959

RESULT 7
US-09-441-411-5
; Sequence 5, Application US/09441411
; Publication No. US2003008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441.411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-5

Alignment Scores:
Pred. No.: 0 Length: 4473
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

```

DB:	11	Gaps:	0
SEQ3 (1-653) x US-09-441-411-5 (1-4473)			
QY	1	MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla	20
DB	175	ATGAGAGTGGCGGCTTGTGCGCTGGGGCTCTCTCTCGCCCTTGTGCCCCCGAGCC	234
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
DB	235	GGAGCACCACCAAGTGTGACCGGCACACATGAAGCTGCGCTCCTGCTCCAGTCCCGAG	294
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrgInGlyCysGlnValValGlnGlyAsnLeu	60
DB	295	ACCCACCTGGACATGCTCGGCACCTTACCAAGGCTGCCAGGTGGTGCAGGGAACCTG	354
QY	61	GluLeuThrTyrProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
DB	355	GAACCTACCTACTCTGCCACCAATGCCAGCTTCTCTCTGCAGGATATCCAGGAGGTG	414
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
DB	415	CAGGCTACGTCTCATGCTCACAACTGAAGTGAAGTCCACTGCAGAGGCTGCGG	474
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
DB	475	ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTAGACAATGA	534
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
DB	535	GACCCGCTGAACAATACCAACCCCTGTACAGGGGCTTCCCAAGGAGGCTTGGGGAGCTG	594
QY	141	GlnLeuArgSerLeuThrGlnLeuLysGlyValLeuIleGlnArgAsnProGln	160
DB	595	CAGCTTGAAGCTCACAGATCTTGAAGAGGGGCTTGTATCCAGGGAACCCCGAG	654
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	180
DB	655	CTCTGCTACCAAGACACGATTTGTGGAAGGACATCTTCCACAAGAACCAACAGCTGGCT	714
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
DB	715	CTCACACTGATAGACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTAAAG	774
QY	201	GlySerArgCysTrpGlyCysSerGluAspCysGlnSerLeuThrArgThrValCys	220
DB	775	GGCTCCCGTCTGGGAGAGTCTTGAGGATTGTGAGGCTTACAGGCTTACGCGCACTGTCTGT	834
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys	240
DB	835	GCCGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGTGCCATGAGCAGTGT	894
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
DB	895	GCTGCCGGTGCACGGGCCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCAACCA	954
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
DB	955	AGTGGATCTGTGAGCTGCACTGGCCAGCCCTGGTCACCTCAACACAGACACGTTTGA	1014
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
DB	1015	TCCATGCCCAATCCCGAGGGCGGTATACATTCGGGGCCAGCTGTGTACTGCTGTCCC	1074
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
DB	1075	TACAACCTACTTTTACGGACGTGGGATCCTCACCCCTGTCTGCCCTGTGCAACACAA	1134
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
DB	1135	GAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAAAGTGCAGCAAGCCCTGTCCCC	1194
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360

DB	1195	GTGTGCTATGTCTGGCATGGAGCAGCTTTCGAGAGTTCGAGGCACTTACCAGTGCCTAAT	1254
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
DB	1255	ATCCAGGAGTTTGTCTGGCTGCAAGAAGATCTTTGGAGAGCTGGCATTTCTGCGGAGAGC	1314
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
DB	1315	TTTGATGGGACCCAGCCCTCCAACTGCCCCGCTCCAGCAGAGCAGCTCCAAAGTGT	1374
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrlleSerAlaTrpProAspSerLeuPro	420
DB	1375	GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCGGACAGCTGCCT	1434
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
DB	1435	GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGACGAATCTTCACAAATGGCGC	1494
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
DB	1495	TACTCGCTGACCTGCAAGGGCTGGCATCAGCTGGCTGGGCTGGCTCCTACTGAGGAA	1554
QY	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
DB	1555	CTGGCAGTGGAGTGGCCCTCATCCACATAACACCCACTCTGCTTCGTGCACACGGTG	1614
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
DB	1615	CCCTGGGACCACTCTTTTGGAAACCCGACCAAGCTCTGCTCCACACTCCCAACCGGCA	1674
QY	501	GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
DB	1675	GAGGACGAGTGTGGGCGAGGGCTGGGCTGCCACCACTGTGCCCGCCGAGGCACTGC	1734
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
DB	1735	TGGGTGCCAGGCCCCACCCAGTGTGTCAACTGCAGCAGTTCCTTCGGGGCCAGGAGTGC	1794
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
DB	1795	GTGGAGGAATGCCAGTACTGCAGGGGCTCCCGCAGGAGTATGTGAATGCCAGGCACTGT	1854
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
DB	1855	TTGCCGTGCCACCCCTGAGTGTGAGCCCGCAGAAATGGCTCAGTACCTGTTTGGACCGAG	1914
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
DB	1915	GCTGACCACTGTGTGGCTGTGCCCACTATAAGGACCCCTCCCTTCTGCGTGGCCGCTGC	1974
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
DB	1975	CCCAGCGTGTGAAACCTGACCTCTCATATGCCCATCTGGAAGTTTCCAGATGAGGAG	2034
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
DB	2035	GGCGCATGCCAGCCCTGGCCCCATCACTGACCCCACTCTCTGTGTGACCTGTGATGACAG	2094
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
DB	2095	GGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGACGTCC	2133

RESULT 8

US-10-101-510-81
; Sequence 81, Application US/10101510
; Publication NO. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510


```
Db 2035 GCGCATGCCAGCCTTGCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAG 2094
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 2095 GGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGAGGTCC 2133

RESULT 9
US-10-146-473-32
; Sequence 32, Application US/10146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Rao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146.473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-473-32

Alignment Scores:
Pred. No.: 0 Length: 4473
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

SEQ3 (1-653) x US-10-146-473-32 (1-4473)
QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGTGGCGGCTGTGCGCTGGGGCTCTCTCGCCCTTGCCCGCGGAGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCCAGCTGGACATGCTCCGCCACTCTACCAGGCTGCCAGGTGGTCAGGGAAACCTG 354
QY 61 GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 355 GAATCTACCTACCTTGCCCAACCAATGCCAGCCTGCTCTCTCTCCAGGATATCCAGGAGTG 414
QY 81 GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGGCTACGTGCTCATGCTCACAACCAAGTAGGAGGAGGCTCCCACTGCAGAGGTGGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTGGCAGGACCCAGCTCTTTGAGGACAACTATGCCCTTGGCCGCTGTAGACAAATGGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGACATACCAACCCCTGTACAGGGGCTCCCGAGGAGGCTGCGGGAGCTG 594
QY 141 GlnLeuArgSerLeuThrGluLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTCAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTTGATCCAGCGGAACCCCGAG 654
```

```
QY 161 LeuCysTyrrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 655 CTCGTCTACAGGACACGATTTTGTGGAGGACATCTTCCACAAAGAACACACGCTGGCT 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 774
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 775 GGCTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTCAGAGGCTCAGCGCATCTCTGT 834
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 835 GCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGTCATGAGGAGGAGCTGT 894
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 895 GCTGCGGCTGCACGGGCCCCAAGCACCTCTGACTGCCCTGGCTGCCCTCCACTTCAACCA 954
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGlu 280
Db 955 AGTGGCATCTGTGAGCTGACTGCCAGCCCTGGTGCACCTACAAACACACAGACAGCTTG 1014
QY 281 SerMetProAsnProGluGlyArgTyrrPheGlyAlaSerCysValThrAlaCysPro 300
Db 1015 TCCATGCCCAATCCGAGGGCCGTTATACATTCGGCGCCAGCTGTGACTGCTCCTGCC 1074
QY 301 TyrAsnTyrrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1075 TACAACCTACCTTTCTACGGAGCTGGATCCTGCACCCCTCGTGGCCCTGCACAACCAA 1134
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1135 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTGACAGAGCCCTGTGCCCA 1194
QY 341 ValCysTyrrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1195 GTGTGCTATGCTTGGGCTGGAGCATTGCGAGAGGTGAGGGCAGTTACCAAGTCCCAAT 1254
QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1255 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGAGCCTTGGCATTTCTTCCGCGAGAG 1314
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1315 TTTGATGGGAGCCAGCCTCCAAACACTGCCCGCTCCAGCCAGAGACAGCTCCAAGTGT 1374
QY 401 GluThrLeuGluGluIleThrGlyTyrrLeuTyrrIleSerAlaTrpProAspSerLeuPro 420
Db 1375 GAGACTCTGGAAGAGATCAGAGTTACCTATATCATCTCAGCATGTCGCGGACAGCCTGCT 1434
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1435 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGAGCAATTTCTGCACAAATGGCGCC 1494
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1495 TACTCGCTGACCCCTGCAAGGGCTGGCATCAGCTGGCTGGGCTCGCTCCTACTCAGGGA 1554
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1555 CTGGCAGTGACTGGCCCTCATCCACCATACACCCACCTCTGCTTCTGTCACACAGGTG 1614
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1615 CCCTGGGACAGCTCTTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCCAACCGGCA 1674
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1675 GAGCAGAGTGTGTGGCGAGGGCTGGCCTGCCACAGCTGTGCCCGCCGAGGAGGACTGC 1734
```



```

1495 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGCTGGGCTGGCTCACTGAGGGAA 1554
461 LeuGlySerGlyLeuAlaLeuLeuHisHisAsnThrHisLeuCysPheValHisThrVal 480
1555 GTGGGAGTGGAGTGGCCCTCATCCACCAATACACCCACTCTGCTGCTGACACGGTG 1614
481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
1615 CCCTGGGACCAGCTCTTCGGAAACCGCACCAGCTCTGCTCCACACTGCCAACCGGCCA 1674
501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
1675 GAGGACGAGTGTGTGGCGAGGGCTGGCTGCTCCACAGCTGTGGCCCGAGGCACTGC 1734
521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
1735 TGGGGTCCAGGGGCCACCCAGTGTGTCAACTGCAGCCAGTGTCTTCGGGGCCAGGAGTGC 1794
541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
1795 GTGGAGGAATGCCGAGTACTGAGGGCTCCCGAGGGAGTATGTGATGCCAGGCACTGT 1854
561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
1855 TTGCGCTGCGACCTGAGTGTGAGCCCAAGATGGCTCAGTACCTGTTTGGACCGGAG 1914
581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
1915 GCTGACCACTGTGTGGCTGTGCCCACTATAAGGACCCCTCTTCTGCGTGGCCGCTGC 1974
601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
1975 CCCAGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2034
621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
2035 GCGCATGCGACGCTTGCCCCATCAACTGCACCCACTCTGTGTGGACCTGGATGACAAG 2094
641 GlyCysProAlaGlnArgAlaSerProLeuThrSer 653
2095 GGCTGCCCCCGCAGCAGAGAGCCAGCCCTCTGACGTCC 2133

```

RESULT 11

```

US-09-877-177-11
; Sequence 11, Application US/09877177
; Publication No. US20020192652A1
; GENERAL INFORMATION:
; APPLICANT: Peter V. Danenberg et al.
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-New Gene Expression
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177-11

```

```

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 3628.00 Matches: 633
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

```

SEQ3 (1-653) x US-09-877-177-11 (1-4530)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20

```

151 ATGGAGCTGGCGCCTTGTGCGCTGGGGCTCTCTCTCGCCCTCTTGGCCCCCGAGCC 210
21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
211 GCGAGCACCACAAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 270
41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
271 ACCCAGCTGACATGCTCCGCCACCTCTACAGGGCTGCAGGTGGTGCAGGGAACCTCG 330
61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
331 GAACTCACCTACTCTGCCACCAATGCCAGCTGTCTCTCTGCGAGGATATCCAGGAGGTG 390
81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
391 CAGGGCTAGCTGCTCATCTGCACCAAGTGAGCAGGTCCACATGCGAGAGCTGCGG 450
101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
451 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCCTGTAGACAATGGA 510
121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
511 GACCCGCTGAACAATACCACTCTGTACAGGGGCTCTCCCGAGGAGGCTGCGGGAGCTG 570
141 GlnLeuArgSerLeuThrGluLeuLysGlyValLeuIleGlnArgAsnProGln 160
571 CAGCTTCGAAGCTCACAGAGATCTTGAAGAGAGGGTCTTGATCCAGGCGGAACCCCGAG 630
161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
631 CTCTGCTACCAGACACGATTTTGTGAAGAGACATCTTCCACAAGAAACCAACAGCTGGCT 690
181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
691 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCGACCCCTGTGTCTCCGATGTGTAAG 750
201 GlySerArgCysTrpGlyCysSerGluAspCysGlnSerLeuThrArgThrValCys 220
751 GGCTCCCGCTGTGGGGAGAGAGTCTTGAGGATTTGTCAGAGCTTGACGGCCTGCTGTGT 810
221 AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
811 GCGGTGGCTGTGCCGCTGCAAGGGCCACTGCCACTGACTGCTGCCATGAGCAGTGT 870
241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
871 GCTGCGGCTGCACGGGCCCCAAGCACTCTGACTGCTGGCTGGCTGCCCTCCACTTCAACCA 930
261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
931 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTGGTACCTACAACACAGACAGCTTGAG 990
281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
991 TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGGCGCAGCTGTGTGACTGCTGCTGCC 1050
301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
1051 TACAACCTACTCTTCTACGGAGCTGGGATCCTGCACCTCTGCTGCGCCCTGCACCAACCA 1110
321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
1111 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1170
341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
1171 GTGTGCTATGCTGTGGCATGGAGCATTGCCAGAGGTGAGGGCAGTTACCACTGCCAAT 1230
361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380

```



```

Db 1231 ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGGCAATTTCTCCCGGAGAGC 1290
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1291 TTTGATGGGACCCAGAGCTCCACACACTGCCCGCTCCAGCCAGAGCAGCTCCAGTGT 1350
Qy 401 GluThrLeuGluGluLeuThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1351 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGCGCGAGCAGCTGCCT 1410
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1411 GACCTCAGCGTCTCCAGAACCTCAAGTAATCCGGGAGCAATTTCTGCACAAATGGCGCC 1470
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1471 TACTCGCTGACCTTGCAAGGCTGGGCATCAGCTGGCTGGGGCTCGCTCAGGGAA 1530
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1531 CTGGCAGTGAGCTGGCGCTCATCCACATAACACCCACCTCTGCTTGCACACAGGTG 1590
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1591 CCCTGGGACCAAGCTCTTTGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1650
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1651 GAGGACGAGTGTGGGGAGGGCTGGCTGCCACAGCTGTGGGCCGAGGCGACTGC 1710
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1711 TGGGGTCAGGGGCCACCCAGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGAGTGC 1770
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAlaAsnAlaArgHisCys 560
Db 1771 GTGAGGAATGCCAGTACTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1830
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1831 TTGGCGTGCCACCTCGAGTGTGAGCCAGCAATGGCTCAGTGACCTGTTTGGACCGAG 1890
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrIleAspProPheCysValAlaArgCys 600
Db 1891 GCTGACCAAGTGTGGCGCTGTGCCCACTATAAGGACCTCCCTTCTGCGTGGCGCGTGC 1950
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1951 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCACTTGAAGTTTCCAGATGAGGAG 2010
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 2011 GGCATGCCAGCTTGCCCATCACTGACCTGACCCACCTCTCTGTGGACCTGGATGACAAG 2070
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 2071 GGCTGCCCGCCGAGAGAGAGCCAGCCCTCTGACGTCC 2109

```

RESULT 12

```

; Sequence 119, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A

```

```

; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: v-erb-b2 avian erythroblastic leukemia viral
; OTHER INFORMATION: oncogene homolog 2 (neuro/glioblastoma derived
; OTHER INFORMATION: oncogene homolog) (ERBB2) gene.
US-10-007-926A-119

```

Alignment Scores:

```

Pred. No.: 0 Length: 4530
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

```

SEQ3 (1-653) x US-10-007-926A-119 (1-4530)

```

Qy 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGGAGCTGGCGGCTGTGGCGGTGGGGGTCTCTCTCGCCCTCTTGCCCGCCGGAGCC 210
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCGAGCAGCCCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 270
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGTCTCCGCCACTCTTACCAGGGCTGCCAGGTGGTGCAGGAAACCTG 330
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAACCTACCTACCTGCCCAATGCCAGCTGCTCTCTCTCCAGGATATCCAGGAGTG 390
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACGTGCTCATCGCTCACAACCAAGTGCAGGAGGTCCCACTGCAGAGGTGCGG 450
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGCGAGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGCTGCTAGACAATGGA 510
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAATACCAACCCCTGTACAGGGGCTCCCGCAGGAGGCTCGCGGAGCTG 570
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCAAGCCCTCAGAGATCTTGAAGAGAGGGGTCTTTGATCCAGCGGAAACCCCGAG 630
Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGTACCAAGGACAGATTTTGTGAAGAGACATCTTCCACAAGACAACACAGCTGGCT 690
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCGCTCTCGGSCCTGCCACCCCTCTTCTCCGATGTGTAAG 750
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGCTGGGAGAGAGTCTCAGGATTTGTGAGAGCTGTCAGAGCCTGACGCGACTGCTGT 810
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 811 GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCTGCTGCTGCTGCTGCT 870

```



```

QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGCTACCAAGGACACGATTGTTGGAAAGACATCTTCCACAAGAACCAACAGCTGGCT 690
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACATGATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGAAG 750
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GSCTCCCGCTCTGGGAGAGAGTTCTGAGAGATTGTGAGACCTTGACGGCAGCTGTCTGT 810
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db 811 GCCGGTGGCTGTGCCCGCTGCAAGGGCCACTGTCCTGCTGCTGCTGCTGCTGCTGCTGCT 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGCCGCTGCACGGGCCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 931 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGCTACCTACCAACACAGACACGTTTGAG 990
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 991 TCCATGCCCAATCCGAGGGCGGTATACATTCCGGCCAGCTGTGTGACTGCTGCTGCTGCC 1050
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1051 TACAACATACCTTTCTACGGAGCTGGGATCTCTGCACCCCTGCTGCTGCTGCTGCTGCTGCT 1110
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1111 GAGCTGACAGCAGAGGATGAACACACGCGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1170
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1171 GTGTGCTATGCTTGGGCATGGAGCACTTCCGAGAGTGAGGGCAGTTACCAAGTGCCTAAT 1230
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1231 ATCCAGGAGTTGCTGCTGCACAGAAGATCTTTGGGAGCTGGCATTTCTGCCGGAGGC 1290
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
Db 1291 TTTGATGGGACCCAGCCTCCAAACACTGCCCGCTCCAGCCAGCAGCAGCTCCAAGTGT 1350
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1351 GAGACTTGAAGAGATCAAGGTTACCTATACATCTCAGCATGGCCGAGCAGCCTGCCT 1410
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyVala 440
Db 1411 GACCTCAGCGCTCTCCAGAACCTGCAGTAATCCGGGGAGCAATCTGCACAATGGCGCC 1470
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1471 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAA 1530
QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1531 CTGGGAGTGGATGGCCCTTCATCCACCAATTAACACCCACCTCTGCTTCTGTCACAGGGT 1590
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1591 CCCTGGGACCACTCTTTTGGAAACCCGACCAAGCTCTGCTCCACTGCTCCACACCGGCCA 1650
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1651 GAGGACAGTGTGTGGGCGAGGGCTTGGCTGCTGCCACAGCTGTGCCCGCGGCGCACTGC 1710

```

```

QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1711 TGGGTTCCAGGGCCACCCAGCTGTCAACTGCAGCAGTTCTTCGGGGCCAGAGTGC 1770
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1771 GTGGAGGAATCCGAGTACTGCAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACGT 1830
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1831 TTGCCGTCGCCACCTGAGTGTGAGCCCGAGAAATGGCTCAGTGACTGTTTGGACCGGAG 1890
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db 1891 GCTCACCAGTGTGTGGCTGTGCCACTATTAAGACCCCTCTCTTCGTCGGCCCGCTGC 1950
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1951 CCCAGCGTGTGAACCTGACCTCTCTACATGATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 2011 GGCGCATGCCAGCTTGCCTCATCACTGCACCCACTCTCTGTGTGGACTGGATGACAAG 2070
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 2071 GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 2109

```

RESULT 14

US-10-338-730-1

; Sequence 1, Application US/10338730

; Publication No. US20030147905A1

; GENERAL INFORMATION:

; APPLICANT: Genzyme Corporation

; APPLICANT: Nicolette, Charles A.

; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS

; FILE REFERENCE: 5017C

; CURRENT APPLICATION NUMBER: US/10/338,730

; CURRENT FILING DATE: 2003-01-08

; PRIOR APPLICATION NUMBER: US 09/527,487

; PRIOR FILING DATE: 2002-03-16

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 4530

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (151)..(3915)

; OTHER INFORMATION:

US-10-338-730-1

Alignment Scores:

Pred. No.: 0 Length: 4530

Score: 3628.00 Matches: 653

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

SEQ3 (1-653) x US-10-338-730-1 (1-4530)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyVala 20

Db 151 ATGAGACTGGCGCCTTGTGCCCTGGGGCTCCTCTCTGCCCTCTTGGCCCCCGGAGCC 210

QY 21 AlaserThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40

Db 211 GCAGCACCAAGTGTGCACCGCAGACATGAAGCTGGGGCTCCCTGCTCCAGTCCCGAG 270

QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60

Db 271 ACCACCTGGACATGCTCGCCACCTCTTACCAGGGCTGCCAGGTGTGCGAGGAACCTG 330
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAACCTACCTACTCGCCACCAATGCCAGCCTGTCTTCTTCGAGGATATCCAGGAGTG 390
QY 81 GlnGlyTyrValLeuIleAlaHisGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACGTCTCATCGCTCACAAACCAAGTGAAGCAGGTCCCACTGCAGAGGCTGCGG 450
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGCGAGGACCCAGCTCTTTGAGGACACTATGCTTGGCCGTGTAGACAAATGGA 510
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAATACCAACCCCTGTTCACAGGGGCTCCCGAGGAGGCTGCGGAGCTG 570
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCGAAGCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGGGGAACCCACAG 630
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db 631 CTCTGCTACAGGACACGATTTTGTGAAGGACATCTTCCACAAGAACCAACCAAGCTGGCT 690
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAG 750
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGTGGGGAGAGAGTCTGAGGATGTGTACAGAGCTGACGGGCATGCTGTGT 810
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db 811 GCCGGTGGCTGTGCCCGCTGCAAGGGGGCCACTGCCCCACTGACTGTGCCATGAGCAGTGT 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGCCGGCTGACGGGCCCCAAGCAGCTCTGACTGCCTGGCTGGCTTCCACTTCAACCCAC 930
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 931 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACAACACAGACACAGTTTGAG 990
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 991 TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCC 1050
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1051 TACAACCTACCTTTCTACGGACGTGGATCTCGACCCCTGCTGCCCTCGCCACCAACAA 1110
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1111 GAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCA 1170
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1171 GTGTGCTATGGTCTGGCATGGAGCATTGCGAGAGGTGAGGGCAGTTACCAAGTCCCAAT 1230
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1231 ATCCAGGAGTTTGTCTGGCTGCAAGAAGATCTTTTGGAGCTTGGCATTTCTGCCGGAGAGC 1290
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1291 TTTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGACAGCTCCAAGTGTT 1350
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1351 GAGACTCTGGAAGAGATCACAGTTTACCTATACATCTCAGCATCTCAGCATGGCCGACAGCTGCT 1410

RESULT 15

US-10-177-293-125

; Sequence 125, Application US/10177293

; Publication No. US20030124128A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Glatt, Karen

; APPLICANT: Zhao, Xumei

; APPLICANT: Gannavarpu, Manjula

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Mertens, Maureen

; APPLICANT: Myer, Vic

; APPLICANT: Wang, Youzhen

; APPLICANT: Xu, Yongyao

; APPLICANT: Hoersch, Sebastian

; APPLICANT: Monahan, John

; APPLICANT: Meyers, Rachel E.

; APPLICANT: Bast Jr., Robert C.

; APPLICANT: Hortobagyi, Gabriel N.

; APPLICANT: Pusztai, Lajos

; APPLICANT: Meric, Funda

; APPLICANT: Sahin, Aysegul

; APPLICANT: Mills, Gordon B.

; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT

QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653

Db 2071 GGCTGCCCGCCGCGAGCAGAGCCAGCCCTCTGACGTCC 2109

QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyYargIleLeuHisAsnGlyAla 440

Db 1411 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCGCACAATGCGGCC 1470

QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460

Db 1471 TACTCGCTCACCCTCGAAGGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGGAA 1530

QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480

Db 1531 CTGGGCGAGTGGACTGGCCCTCATCCACATAACACCCACTCTGTCTTCTGTGCACACGGTG 1590

QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500

Db 1591 CCCTGGGACCACTCTTTCGGAACCGCCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1650

QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520

Db 1651 GAGGACGAGTGTGTGGCGAGGGCTTGGCTGCGCACAGCTGTGCGCCGAGGCGCACTGC 1710

QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540

Db 1711 TGGGGTCCAGGGCCCAACCAAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1770

QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560

Db 1771 GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGTAATGCCAGCACTGT 1830

QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580

Db 1831 TTGCGCTGCGACCTGAGTGTGAGCCCAAGAGTGGCTGAGTACCTGTTTGGACCGGAG 1890

QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600

Db 1891 GCTGACCAAGTGTGGCCCTGTGCCACTATAAGGACCCCTCCCTCTGCGTGGCCCGCTGC 1950

QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620

Db 1951 CCCAGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAAGTTTCCAGATGAGGAG 2010

QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640

Db 2011 GGCGCATGCCAGCTTGCCCATCAACTGCACCACTCCTCTGTGTGACCTGGATGACAAG 2070

; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
 ; FILE REFERENCE: MRI-038
 ; CURRENT APPLICATION NUMBER: US/10/177,293
 ; CURRENT FILING DATE: 2002-06-21
 ; PRIOR APPLICATION NUMBER: US 60/299,887
 ; PRIOR FILING DATE: 2001-06-21
 ; PRIOR APPLICATION NUMBER: US 60/301,572
 ; PRIOR FILING DATE: 2001-06-27
 ; PRIOR APPLICATION NUMBER: US 60/306,501
 ; PRIOR FILING DATE: 2001-07-18
 ; PRIOR APPLICATION NUMBER: US 60/325,002
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US 60/362,585
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
 ; PRIOR FILING DATE: 2002-05-14
 ; NUMBER OF SEQ ID NOS: 506
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 125
 ; LENGTH: 4530
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-177-293-125

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
0	3628.00	4530	553
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

SEQ3 (1-653) x US-10-177-293-125 (1-4530)

Qy	1	MetGluLeuAlaLeuCysArgTyrGlyLeuLeuAlaLeuLeuProGlyVala	20
Db	151	ATGGAGCTGGCGGCTGTGCGCTGGGGGCTCTCTCGCCCTTGTCCCGGAGGCC	210
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	211	CGAGACCCCAAGTGTGCACCGGCACACATGAAGCTGCGCTCCCTGCCAGTCCCGAG	270
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
Db	271	ACCCACCTGGACATGCTCCGGCCACTCTTACCAGGGCTGCCAGGTGGTGCAGGGAACCTG	330
Qy	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	331	GAACCTACCTACCTGCCCAATGCCAGCTGTCTTCTCGAGGATATCCAGAGGTG	390
Qy	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	391	CAGGGCTACGTGCTCATCGCTCACCAACCAAGTAGGCAGGCTCCCACTGCAGAGCTGGG	450
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	451	ATTGTGGAGGACCCAGCTCTTGTGAGSACAACCTATGCCCTGGCGCTGCTAGACAATGGA	510
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	511	GACCGCTGACACATACCCCTGTGCACAGGGGCTCCCGCAGAGGCTCGCGGGAGCTG	570
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Db	571	CAGCTTGAAGCTCAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCGAG	630
Qy	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	631	CTCTGCTACAGGACACGATTTGTGAAGGACATCTCCACAAGAACCAACACAGCTGGCT	690
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	691	CTACACTGATAGACCAACCCCTCTCGGGGCTTGCACCCCTGTCTTCTCCGATGTGAAG	750

Qy	201	GlySerArgCysTyrGlySerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	751	GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATGTCAGAGCTGACGCGACTGTCTGT	810
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	811	GCCGGTGGCTGTGCCCGCTGCAAGGGGCACTGCCCACTGACTGCTGCATGAGCAGTGT	870
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	871	GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGGCTGCTCCTCCTCCTCAACAC	930
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	931	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTGCCTACACACAGACACAGTTGAG	990
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	991	TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGCTGCTGCTGCC	1050
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisGln	320
Db	1051	TACAACTACCTTCTACGGACGTGGATCCTGCACCCCTGCTGCTGCCCTGCAACACAA	1110
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1111	GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAGAGTGCAGCAACCCCTGTGCCGA	1170
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1171	GTGTGCTATGCTTGGCATGAGCAGCTTGCAGAGGTGAGGGCAGTTACCACTGCCAAT	1230
Qy	361	IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1231	ATCCAGGAGTTGTGCTGCAAGAGATCTTGGAGCCTGGCATTTCTGCCGGAGAGC	1290
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1291	TTTGATGGGACCCAGCCTCCACACACTGCCCGCTGCCAGCCAGACAGCTCCCAAGTGT	1350
Qy	401	GluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro	420
Db	1351	GAGACTCGAAGAGATCACAGTTACCTATACATCTCAGCATGCCCGCAGACAGCTGCT	1410
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisGlnGlyAla	440
Db	1411	GACCTCAGCGTCTTCCAGAACCTCAAGTAATCCGGGACGAATTCGCAATGGCGCC	1470
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1471	TACTCGCTGACCTGCAAGGCTGGCATCAGCTGGCTGGGCTGCGCTCCTCCTGAGGAA	1530
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1531	CTGGCAGTGGACTGGCCCTCATCCACATAACACCCACCTCTGCTGCTGCACACAGGTG	1590
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1591	CCCTGGGACAGCTCTTTCGGAAACCCGACCACTGCTGCTCCACACTGCCAACCGGCCA	1650
Qy	501	GluAspGlyCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1651	GAGGACAGGTGTGGGGCAGGGGCTGGCTGCCCTGCCAGCTGTCGGCCGAGGACACTGC	1710
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1711	TGGGTCAGGGCCCAACCCAGTGTCAACTGCAGCAGTCTCTTCTTCGGGGCCAGGAGTGC	1770
Qy	541	ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1771	GTGGAGGAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1830

Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1831	TTCCGCTGCCACCCCTGAGTGTGACGCCCAAGATGGCTACGTGACCTGTTTGTGACCGGAG	1890
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAlaGqCys	600
Db	1891	GCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCCTCCCTCTCGGTGGCCCGCTGC	1950
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu	620
Db	1951	CCGACGGGTGAACCTGACCTCTCTCATACGCCATCTGGAAAGTTTCCAGATGAGGAG	2010
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2011	GGCGCATGCCAGCCTTGCCCCCATCACTGCACCCCACTCCCTGTGTGGACCTGGATGACAAG	2070
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2071	GGGTGCCCGCCGAGCAGAGAGCGACGCTCTCACGTCC	2109

Search completed: October 17, 2003, 03:54:11
Job time : 732.009 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 620.221 Seconds
(without alignments)
2842.104 Million cell updates/sec

Title: SEQ3
Perfect score: 3628
Sequence: 1 MELAALCRWGLLLALLPPGA.....CVDLDDKGPAPQAPASPLTS 653

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-xlp
-MODEL-frame_p2n.model -DEV-xlp

-Q/cgn2_l/USPTO.spool_p/HOLLERAM480/runat_15102003_131912_20526/app_query.fasta_1.4685
-DB-N.Geneseq_19Jun03 -OFMT-fastap -SUFFIX-rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT=ptc -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER-HOLLERAM480.qcgn_1.1.1758 -runat_15102003_131912_20526 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3628	100.0	2385	18	AAT72725 Her-2-GM-CSF immuno
2	3628	100.0	3600	21	AAH89736 Human HER-2/neu co
3	3628	100.0	3768	17	AAT40739 HER-2/neu oncogene
4	3628	100.0	3768	20	AAH01912 Human HER-2/neu on
5	3628	100.0	3768	21	AAH09455 Human heregulin 2
6	3628	100.0	3768	22	AAH23392 Human HER-2/neu pr
7	3628	100.0	3768	24	ABZ35744 Human ERBB2 polyu
8	3628	100.0	3768	24	ABX0987 Human ERBB2 DNA fr
9	3628	100.0	3768	24	AAH43935 Human HER-2 CDNA.
10	3628	100.0	3768	24	AAH43986 Human ERBB2 antigen
11	3628	100.0	3768	24	ABV78168 Human ERBB2 DNA SE
12	3628	100.0	3768	24	AAH32743 Human Her-2/neu pr
13	3628	100.0	3768	24	ABA92250 Human Her-2/neu CD
14	3628	100.0	3768	24	ABK10730 Human Her-2/neu DN
15	3628	100.0	3768	24	ABL91709 Human polynucleoti
16	3628	100.0	3768	24	ABK14058 Human HER2 (Erbb2)
17	3628	100.0	4472	21	AAA14812 cDNA encoding the
18	3628	100.0	4473	19	ABQ76220 Human tumour anti
19	3628	100.0	4473	20	AAZ31071 HER-2 nucleic acid
20	3628	100.0	4473	24	ABZ34969 Human gene express
21	3628	100.0	4473	24	AAH38904 Human Her-2 DNA.
22	3628	100.0	4530	16	AAT01585 Her-2/neu (ERBB2/c
23	3628	100.0	4530	18	AAT71253 Human HER2 gene.
24	3628	100.0	4530	21	AAZ60815 Nucleotide sequenc
25	3628	100.0	4530	22	AAD19731 Human-tyrosine kin
26	3628	100.0	4530	24	ABZ35012 Human gene express
27	3628	100.0	4530	24	ABV94128 Breast carcinoma r
28	3628	100.0	4530	24	ABN85585 Human HER2-neu SEQ
29	3628	100.0	4530	24	ABK83918 Human CDNA differe
30	3628	100.0	4530	25	ACC50139 Breast cancer asso
31	3628	100.0	4530	25	ABQ83856 Human Her2/Neu enc
32	3628	100.0	9274	22	AAF24297 HER2 transgene pla
33	3628	100.0	9274	24	AAH43934 HER-2 transgene pl
34	3628	100.0	9274	24	ABK14057 Human HER2 (Erbb2)
35	3585	98.8	4299	14	AAZ45086 Sequence encoding
36	3525	97.2	2871	21	AAZ50586 DC8scfv-erbB2EC fu
37	3458	95.3	3678	24	ABK86207 cDNA encoding huma
38	3422	94.3	1872	11	AAQ06828 Extracellular port
39	3106	85.6	3955	16	AAT01590 Rat neu promoter.
40	3106	85.6	3955	21	AAA89753 Mouse Her-2/neu prot
41	3104.5	85.6	2763	24	ABA92252 Her-2/neu extracel
42	3104.5	85.6	2781	24	ABA92253 Mouse Her-2/neu CD
43	3100.5	85.5	3771	21	AAA89737 Nucleotide sequenc
44	3100.5	85.5	3771	22	AAH42210 Mouse Her-2/neu CD
45	3100.5	85.5	3771	24	ABA92251

ALIGNMENTS

RESULT 1
AAT72725
ID AAT72725 standard; cDNA: 2385 BP.
XX
AC AAT72725;
XX
DT 17-SEP-1997 (first entry)
XX
DE Her2-GM-CSF immunostimulant fusion protein DNA.
XX
KW Her2-GM-CSF; granulocyte macrophage colony stimulating factor;
KW growth factor receptor; oncogene; immunostimulant; cancer;
KW therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers


```

FT CDS 11..2359
FT /*tag= a
FT /product= GM-CSF-Her2 fusion protein
FT mRNA 11..1969
FT /*tag= b
FT /product= Her2
FT mRNA 1970..1975
FT /*tag= c
FT /product= Leu-Glu linker
FT mRNA 1976..2359
FT /*tag= d
FT /product= GM-CSF
XX
XX WO9724438-A1.
XX
XX 10-JUL-1997.
XX
XX 23-DEC-1996; 96WO-US20241.
XX
XX 28-DEC-1995; 95US-0579823.
XX
XX (ACTI-) ACTIVATED CELL THERAPY INC.
XX
XX Laus R, Ruegg CL, Wu H;
XX
XX WPI; 1997-363674/33.
XX P-PSDB; AAW19764.
XX
XX Potent APC that activates T-cells to give multivalent cellular
XX immune response - can also induce a cytotoxic T-cell response in a
XX vertebrate subject
XX
XX Disclosure; Fig 8; 45pp; English.
XX
XX A nucleic acid molecule (AAT72725) codes for a fusion protein
XX (AAW19764) comprising granulocyte-macrophage colony stimulating
XX factor (GM-CSF) and Her2, a growth factor receptor that is
XX over-expressed in breast and ovarian cancer cells. It was
XX prepd. by PCR amplification of Her2 cDNA from a breast cancer
XX cell line and fusion to GM-CSF cDNA. Fusion expression vectors can
XX be used to transfect mammalian and insect cells. The Her2-GM-CSF
XX fusion protein is used to generate anti-Her2 immunity. Tumour
XX cells are eliminated by cytotoxic T lymphocytes activated in vivo
XX or in vitro by exposure to antigen-presenting cells exposed to the
XX fusion protein.
XX
XX Sequence 2385 BP; 488 A; 780 C; 677 G; 440 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 8.69e-230 Length: 2385
XX Score: 3628.00 Matches: 653
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 18 Gaps: 0
XX
XX SEQ3 (1-653) x AAT72725 (1-2385)
XX
XX QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
XX DB 11 ATGGAGCTGGCGGCTTGTGCCGCTGGGGCTCTCTCTGCGCCCTTCTTGGCCCGGAGCC 70
XX
XX QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
XX DB 71 GCGAGCACCACAAAGTGTGCACCGGCACACATGAGCTGGCTCCCTGCCAGTCCCGAG 130
XX
XX QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
XX DB 131 ACCACCTGGACATGCTCCGCCACCTCTACCAGGCTGCCAGGTGGTGCAGGGAACCTG 190
XX
XX QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
XX DB 191 GAACCTACCTACCTGCCACCAATGCCAGGCTGTCTTCTTGCAGGATATCCAGGAGGTG 250

```

```

QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
DB 251 CAGGGCTACGTGCTCATCGCTCACAACTGAGGAGGAGTCCCACTGCAGAGCTGGG 310
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
DB 311 ATTTGCGGAGCACCCAGCTCTTTGAGGACAACATATGCCCTGGCCGTGCTAGACAATGA 370
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
DB 371 GACCCGCTGAACAATACCCCTCTCACAGGGGCTCCCAAGAGAGCTGCGGGAGCTG 430
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
DB 431 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 490
QY 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
DB 491 CTCTGCTACCAAGGACACGATTTTGTGAAGGACATCTCCACAAGAACAACCACTGGCT 550
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
DB 551 CTCACACTGATAGACACCAACCGCTCTCGGGGCTGCCACCCCTGTTCTCCGATGTGAAG 610
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
DB 611 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGAGTTGTCAGAGCTGACGCGCACTGCTGT 670
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
DB 671 GCCGGTGGCTGTGCCGCTGCAAGGGGCCACTGCCCACTGACTGCTCCATGAGCAGTGT 730
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 731 GCTGCCGGCTGCAGGGGCCCAAGCACTCTGACTGCTGCGCTGCCCTCCACTTCAACAC 790
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
DB 791 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACAGACACAGCTTTGAG 850
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
DB 851 TCCATGCCCAATCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTGCC 910
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
DB 911 TACAACCTACCTTTCTACGGACGTGGGATCTGCACCTCTGCTGCCCTCCGACCAACAA 970
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 971 GAGGTGACAGCAGAGGATGGAACACACACGGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1030
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1031 GTGTGCTATGTCTGGGCTGAGCAGCTTGCAGAGGTGAGGGCAGTTACAGTGCCTCAAT 1090
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
DB 1091 ATCCAGAGTTGCTGGCTGCAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGAGAGC 1150
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
DB 1151 TTTGATGGGACCCAGCCTCCAACTGCCCGCTCCAGCCAGCAGAGCTCCAGTGT 1210
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
DB 1211 GAGACTTGAAGAGATCAGAGGTACCTATACATCTCAGCATGGCGGAGAGCTGCCT 1270
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
DB 1271 GACCTACGCTCTTCCAGAACCCTGCAAGTAAATCGGGGAGCAATCTTGCAACATGGCGCC 1330

```

```

QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTIPLeuGlyLeuArgSerLeuArgGlu 460
Db 1331 TACTCGTCAACCTCAAGGGCTGGCATCAGCTGGCTGGGCTCACTGAGGGA 1390
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1391 CTGGGACAGTGGACTGGCCCTCATCACCATAACACCCACCTCTGCTTGGTGCACAGGTG 1450
QY 481 ProTIPAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1451 CCTGGGACCACCTCTTTCGGAAACCGCCACCAAGCTCTGCTCCACACTGCCAACCGGCA 1510
QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1511 GAGGACGAGTGTGGGCGAGGCTGGCTGCCACCACTGTGGCCGAGGCACTGC 1570
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1571 TGGGTTCCAGGCCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGCGCAGGAGTGC 1630
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1631 GTGGAGGATGCGGAGTACTGAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1690
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1691 TTGCGCTGCCACCTCAGTGTGAGGCCCAAGATGGCTCAGTCACCTGTTTGGACCGGAG 1750
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db 1751 GCTGACCACTGTGTGGCTGTGCCCACTATAAGSACCCCTCCCTTCTGCTGGCCGCTGC 1810
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu 620
Db 1811 CCCAGGCTGTGAACCTGACCTCTCTACATGCCCACTGGAAGTTTCCAGATGAGGAG 1870
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 1871 GCGCATGCGCAGCTTGGCCCATCAACTGCACCCCACTCTGTGTGGACCTGGATGACAAG 1930
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1931 GGCTGCCCGCGGAGCAGAGCCAGCCCTCTGTGACGTCC 1969

RESULT 2
AAA89736
ID AAA89736 standard; DNA; 3600 BP.
XX
AC AAA89736;
XX
DT 12-JAN-2001 (first entry)
XX
DE Human HER-2/neu coding sequence.
XX
KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
KW colon cancer; ds.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT
FT Location/Qualifiers
FT 1..3600
FT /*tag= a
FT /product= "HER-2/neu protein"
XX
PN WO200044899-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-US02164.
XX
PR 29-JAN-1999; 99US-0117976.
XX
XX

```

```

PA (CORI-) CORIXA CORP.
PA (SMIK ) SMITHKLINE BEECHAM.
XX
PI Cheever MA, Gheysen D;
XX
DR WPI; 2000-505976/45.
DR P-PSDB; AAB211198, AAB21208.
XX
PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
PS Disclosure; Fig 15; 128pp; English.
XX
CC The present sequence encodes the human HER-2/neu protein. HER-2/neu is
CC a member of the tyrosine kinase family of receptor-like glycoproteins
CC and shows homology to the epidermal growth factor receptor (EGFR). It
CC probably plays a part in cell growth and/or differentiation. The
CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising
CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
XX
SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 other;

Alignment Scores:
Pred. No.: 1.43e-229 Length: 3600
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

SEQ3 (1-653) x AAA89736 (1-3600)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProGlyAla 20
Db 1 ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCTCTCGCCCTCTTGGCCCCCGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGACCCCAAGTGTGCACCGGCACACATGAGCTGCGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCCAGCTGGACATGCTCGCCACCTCTACCAAGGCTGCCAGGTGGTGCAGGAAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 181 GAATCAGCTACCTGCCCAATGCCAGCTGTCTTCTCTGCAGGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGCTAGTGTCTATCGCTCACAAACCAAGTGCAGGAGTCCCTCCCTGCAGAGGTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTTGGCAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCGTGTAGACAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCGCTGAACAATACCACCCCTGTACAGGGGCTCTCCCAAGGAGGCTCGGGAGGTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGTTCGAAGCCTCACAGAGATCTTGAAGAGGAGGGTCTTGATCCAGCGGAACCCCGAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCGTCTACAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACCAACAGCTGCT 540

```


CC normal gene product. Nucleotides 2026-3765 of the cDNA sequence
 CC code for the intracellular domain (Lys676-Val1255) of the HER-2/neu
 CC protein, which is useful for immunisation against malignancy.
 CC Nucleic acids can be used to direct expression of the intracellular
 CC domain in transformed host cells, or are used, alone or in a viral
 CC vector, for genetic immunisation of an animal.

XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:

Pred. No.: 1 51e-229 Length: 3768
 Score: 3628.00 Matches: 653
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 17 Gaps: 0

SE03 (1-653) x AAT40739 (1-3768)

QY 1 MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 Db 1 ATGGAGCTGGCGCTTGTGGCGCTGGGGCTCCTCCTCGCCCTCTTGGCCCCCGAGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 61 GCGAGCACCACAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCTGCCAGTCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 121 ACCCACTGGAGATGTCGCCACCTCTACCAAGGCTGCCAGGTGGTGCAGGGAACCTG 180
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 181 GAATCACCCTACTGCCCCACCAATGCCAGCTGTCTTCTCGAGGATATCCAGAGGTG 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 241 CAGGGCTACGTGCTATCCTCACAACCAAGTGAGCGAGTCCCCTGCAGAGGTGCGG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGTGTGTAGACAATGGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 361 GACCCGCTGAACAATACCACCTGTACAGAGGGGCTCCCCAGGAGGCTGCGGAGCTG 420
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 Db 421 CAGCTTCGAAGCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCAG 480
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
 Db 481 CTCTGCTACAGACACGATTGTGGAGGACATCTTCCACAAGAACCAACCAAGCTGGCT 540
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 541 CTCACACTGATAGACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAA 600
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 601 GGCTCCCGCTGTGGGAGAGAGTCTGAGGATTGTGAGGATGTCAGAGCTGCACGCGCACTGCTGT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys 240
 Db 661 GCGGTGGCTGTGCCCGTGCAGGGGGCACTGCCCACTGACTGCTGCTGATGAGCAGTGT 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 721 GCTGCGGCTGCACGGGCCCCAAGCACTCTGACTGCTGGCTGCTCCACTTCAACCCAC 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 1861 GGCGCATGCCAGCCTTGCCCCCATCACTGCACCCACTCCTGTGTGGACCTGGATGACAAG 1920

Db 781 AGTGCATCTGTGAGCTGCCTGCCAGCCCTGGTCACTACACACACAGACACGTTTGAG 840
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 841 TCCATGCCAATCCCGAGGCGGTATACATTTCGGCGCAGCTGTGTACTGCTGCTGCC 900
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 901 TACAACCTACCTTTCTACGGACGTGGATCTCTCACCTCGTCTGCCCTGCACACCAA 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 961 GAGGTGACAGCAGAGGATGGAAACACACGCGTGTGAGAGTGCAGCAAGCCCTGTGCCGA 1020
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1021 GTGTCTATGTCTGGGATGGAGCATTGCGAGAGGTGAGGCGAGTTTACCAGTGCCAAAT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 Db 1081 ATCCAGGAGTTTGTCTGGTGCAGAAGATCTTTGGAGCTTGGCATTTCTGCCGAGAGC 1140
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400
 Db 1141 TTTGATGGGACCCACCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1200
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro 420
 Db 1201 GAGACTCTGGAAGAGATCACAGTTTACCTTATACATCTCAGCATGCGCGAGACCTGCT 1260
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1261 GACCTCAGCTCTTCCAGAACCTGCAAGTAATCCGGGACGAATTTCTGCACAAATGGCGC 1320
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1321 TACTCGCTGACCTCAAGGGCTGGGCATCAGCTGCTGGGCTGCTGCTACTGAGGAA 1380
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1381 CTGGCAGTGGACTGGCCCTCATCACCATAAACCCACCTCTGCTGTGTCACACGGTG 1440
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1441 CCCTGGAGCAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGCCA 1500
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 Db 1501 GAGGACGAGTGTGTGGCGAGGGCTGGCCTGCCACCTGCTGGCGCCGAGGCGACTGC 1560
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 Db 1561 TGGGTTCAGGGCCCAACCTGTCAGTGTCAACTCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 Db 1621 GTGGAGGAATGCCAGTACTGCAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACCTG 1680
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1681 TTGCGTGCACCTCTGAGTGTGAGGCGCCAGAAATGGCTCAGTGACCTGTTTGGACCGG 1740
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 Db 1741 GCTGACCAAGTGTGGCTGTGCCCTATAGAGACCTTCCCTCTGCTGGCGCCGCTGC 1800
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 Db 1801 CCCAGCGTGTGAACCTTGACCTCTCTTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
 Db 1861 GGCGCATGCCAGCCTTGCCCCCATCACTGCACCCACTCCTGTGTGGACCTGGATGACAAG 1920

QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 DB 1921 GGTCGCCCGCGAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 4

AAX01912
 ID AAX01912 standard; DNA; 3768 BP.

XX AC

XX AC

XX AC

XX 21-APR-1999 (first entry)

XX Human HER-2/neu oncogene DNA.

XX DE

XX HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;

XX KW malignancy; treatment; tumour; ss.

XX OS Homo sapiens.

XX XX

XX Key

XX Location/Qualifiers

XX 1..3768

XX /*tag= a

XX /product= "HER-2/neu"

XX /note= "oncogene"

XX 2026..3765

XX /*tag= b

XX /note= "region which elicits immune response"

XX misc_feature

XX US5869445-A.

XX PN

XX 09-FEB-1999.

XX PD

XX 01-APR-1996; 96US-0625101.

XX PF

XX 01-APR-1996; 96US-0625101.

XX PR

XX 17-MAR-1993; 93US-0033644.

XX PR

XX 12-AUG-1993; 93US-0106112.

XX PR

XX 31-MAR-1995; 95US-0414417.

XX XX

XX (UNIW) UNIV WASHINGTON.

XX XX

XX Cheever WA, Disis ML;

XX PI

XX WPI; 1999-152835/13.

XX DR

XX P-PSDB; AAW92406.

XX XX

XX Use of HER-2/neu polypeptides - for eliciting an immune response to

XX an HER-2/neu associated malignancy, particularly for treating or

XX preventing tumours

XX Claim la; Column 23-32; 26pp; English.

XX XX

XX This sequence encodes the human HER-2/neu oncogene protein. A fragment

XX of this protein is used in a method for eliciting or enhancing an immune

XX response to HER-2/neu protein. The polypeptide can stimulate T cells and

XX B cells to produce an immune response to the HER-2/neu protein. The

XX method can be used for immunisation against a malignancy in which the

XX HER-2/neu oncogene is associated and in the treatment of an existing

XX tumour, or to prevent tumour occurrence or reoccurrence.

XX XX

XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 1..51e-229 Length: 3768

XX Score: 3628.00 Matches: 653

XX Percent Similarity: 100.00% Conservative: 0

XX Best Local Similarity: 100.00% Mismatches: 0

XX Query Match: 100.00% Indels: 0

XX DB: 20 Gaps: 0

XX SEQ3 (1-653) x AAX01912 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
 DB 1 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCTCGCCCTTGTTCGCCCGGAGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 61 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTGCCAGTCCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 DB 121 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAACCTG 180
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
 DB 181 GAACCTACCTACCTGCCCAATGCCAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 241 CAGGGCTACGTGCTCATCGCTCACACCAAGTGGCAGGTCCCTCCACTGCAGAGGCTCGG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 301 ATTGTGCGAGGACCCAGCTCTTTGAGACAACTATGCCCTGGCGTGGCTAGACAATGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB 361 GACCGCTGAACAATACCACTCCCTGTCTCACAGGGGCTTCCCGAGGAGGCTCGGAGCTG 420
 QY 141 GluLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
 DB 421 CAGCTTGAAGCCTTCAGAGATCTTGAAGAGGGGTCTTGATCCAGCGAACCCTCCAG 480
 QY 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 DB 481 CTCTGTACCCAGGACAGATTCTTGGAGAGACATCTTCACAGAACACACAGCTGGCT 540
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGGCTGCCACCCCTGTCTCCGATGTGAAG 600
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 601 GGCTCCCGCTGCTGGGAGAGAGTCTGAGGATTGTGAGAGCTTCAGAGCTTCAGCGCCTGTGT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 DB 661 GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCTGCTGCTGCTGT 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 721 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTTACACACAGACAGACGTTTGAG 840
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 841 TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGGCCCGCAGCTGTGTGCTGCTGCTGCT 900
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 901 TACAACCTACCTTCTAGCGGACGTGGGATCTGCACCCCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 961 GAGGTGCACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1020
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1021 GTGTGTATGGTCTGGGCATGGAGCACTTCGAGAGGTGAGGCGAGTTTACAGTGGCAAT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380


```

Db 301 ATTTGGCGAGCCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGCTGAGACAATGGA 360
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGCTGAACAATACCAACCCCTGTACAGGGGCCCTCCAGGAGGCTCGGGAGCTG 420
Qy 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGln 160
Db 421 CAGCTTGAAGCCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 480
Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCGTCTACCAAGGACAGATTTTGTGAAGAGCATCTTCCACAAGAACCAACCACTGGCT 540
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGAAG 600
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTCTCGGGAGAGAGTTCTGAGGATTTGTGAGAGCCTGACGCGCACTGTCTGT 660
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTCGCGGCTGCACGGGCCCAACGACTCTGACTGCTGCCGTGCCCTCCACTTCAACCAAC 780
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGCTGACCTACAAACACAGACACGCTTTGAG 840
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCGTGCC 900
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACATACCTTTCTACGGAGCTGGGATCTGACACCTCTGCTGCCCTCCCTGCACAAACAA 960
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1020
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1021 GTGTGCTATGCTTGGGCATGGAGCACTTGGAGAGGTGAGGGCAGTTACCAAGTCCCAAT 1080
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 ATCCAGAGTTGTGCTGCAAGAAGATCTTTGGGAGCTGGCATTTCTGCCGGAGAGC 1140
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1141 TTTGATGGGACCCAGCCTCCAAACACTGCCCGCTCCAGCCAGCAGACAGCTCCAAGTCTTT 1200
Qy 401 GlnThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1201 GAGACTCTGGAAGAGATCAGAGTTACTATACATCTCAGCATGCGCGGACAGCTGCCT 1260
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTCAGCGTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACATGGCGCC 1320
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGTGCCTTGCAGGGCTGGGCATCAGCTGGCTGGGGCTGCCTCACTGAGGAA 1380
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480

```

```

Db 1381 CTGGGCGAGTGACTGGCCCTCATCCACCATAAACACCACCTCTCTCTGTCGCACACGGTG 1440
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCCTGGGACCAAGCTCTTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGACGAGTGTGTGGCGGAGGGCTGGCCCTGCCACCAAGTGTGCGCCCGAGGCACTGC 1560
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGGTCCAGGGCCCAACCCAGTGTCACTGACGAGCAGTTCTTCGGGGCCAGGATGC 1620
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCAGTACTGCCAGGGCTCCCCAGGAGATGTGTAATGCCAGGCACTGT 1680
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCGGTGCCACCCCTGAGTGTGAGCCCGCAGAAATGGCTCAGTGACCTGTTTGGACCGAG 1740
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1741 GCTGACGAGTGTGGGCTGTGCCCACTATAAGGACCCCTCCCTTCTGGTGGCCGCTGC 1800
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GGCACATGCCAGCCTTGCCCATCACTGACCCACCTCTCTGTGTGGACCTGGATGACAAG 1920
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCCGAGCAGAGAGCCCTCTGACGTCC 1959

RESULT 6
AAH23392
ID AAH23392 standard; DNA; 3768 BP.
XX
AC AAH23392;
XX
XX 25-SEP-2001 (first entry)
XX Human HER-2/neu protein encoding DNA.
XX
XX Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
XX oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..3768
XX FT /*tag= a
XX FT /product= "HER-2/neu protein"
XX
XX WO200153463-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US01850.
XX
XX 21-JAN-2000; 2000US-0177545.
XX
XX (CORI-) CORIXA CORP.
XX
XX Cheever MA, Hand-Zimmermann S;
XX
XX WPI; 2001-476112/51.
XX
XX P-PSDB; AAB85458.
XX

```


PT New antigen-presenting cells, useful as vaccines for eliciting or
 PT enhancing an immune response to HER-2/neu protein, particularly useful
 PT for treating or preventing cancer, e.g. breast cancer
 XX
 PS Claim 1: Page 41-46; 49pp; English.

XX The invention provides an isolated antigen-presenting cell, which
 CC expresses at least an immunogenic portion of a polypeptide that produces
 CC an immune response to HER-2/neu protein. The antigen-presenting cells are
 CC useful as vaccines for eliciting or enhancing an immune response to
 CC HER-2/neu protein, particularly in treating or preventing malignancies in
 CC which the HER-2/neu oncogene is associated. Specifically, these are
 CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,
 CC colon, lung or prostate cancers. The present sequence represents a DNA
 CC encoding the human HER-2/neu protein (also known as p185 or c-erbB2).

XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:

Pred. No.: 1.51e-229 Length: 3768
 Score: 3628.00 Matches: 653
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

SEQ3 (1-653) x AAH23392 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 DB 1 ATGGAGCTGGCGCTTGTGGCGCTGGGGCTTCCTCGCCCTTGTGCCCCCGAGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 61 GCGAGACCCAAAGTGTGACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 DB 121 ACCCACTGGACATGCTCGGCACCTTACCAAGGCTGCCAGGTGGTGCAGGGAACCTG 180
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
 DB 181 GAACATCACTACTCCCAACCAATGCCAGCTGTCTTCTCGAGATATCCAGAGGTG 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 241 CAGGCTAGTGTCTATCGCTCACAAACCAAGTACAGGCTGCCACTGCAGAGGTGCGG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 301 ATTGTGCGAGGCACCAAGCTTTGAGGACAACTATGCCCTGGCGGTGCTAGACAATGGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB 361 GACCCGCTGAACAATACCACTCTACAGGGCTCCCAAGAGGCTTCCAGAGGCTGCGGAGCTG 420
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
 DB 421 CAGCTTCGAGGCTCACAGATCTTGAAGAGGGGTCTTATCCAGCGGAACCCCGG 480
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 DB 481 CTCTCTACCAAGACACGATTTTGGAGGACATCTTCCACAAGAACACACCACTGGCT 540
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 541 CTCACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGCTCCGATGTGAAG 600
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 601 GGCTCCCGCTGTGGGGAGAGATTCTGAGGATTGTGAGGCTTACAGGCTTACGCGCACTCTGT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGluGlnCys 240

DB 661 GCCGCTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 721 GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCGCTGCCCTCCACTTCAACCA 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 781 AGTGCACTGTGAGCTGCACCTGCCCGCTGGTCACTACCTACACACAGACACAGTTTGA 840
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 841 TCCATGCCCAATCCCGAGGCGGTATACATTGCGCGCAGCTGTGTGACTGCTGCTGCC 900
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 901 TACAATACCTTTCTACGAGCGTGGATCTCTACACCTCTGCTGCCCGCTGCACAACCA 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 961 GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAGTGAGAGCCCTGTGCCGA 1020
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1021 GTGTGCTATGTCTGGCATGGAGCACTTGGAGAGGTGAGGCACTTACCACTGCCAAT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1081 ATCCAGAGGTTTGTGCTGGCTGCAAGAAGATCTTTGGGAGCTGGCATTTCTGCCGAG 1140
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB 1141 TTTGATGGGACCCAGCCTCCCAACTGCCCGCTCCAGCCAGAGCAGCTCCAGTGTGT 1200
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 DB 1201 GAGACTCTGGAAGAGATCACAGTTTACCTATACATCTCAGCATGCGCGACAGCTGCT 1260
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1261 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGACGAAATCTGCACAATGCGGC 1320
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1321 TACTGCTGACCTTCAAGGGCTGGCATCAGCTGGCTGGGCTGCGCTCCTCAGGGAA 1380
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1381 CTGGCAGTGGACTGGCTCTCATCCACCAACACCCACCTCTGCTTGTGCACAGGTG 1440
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1441 CCCTGGGACCACTCTTTCCGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCA 1500
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1501 GAGGACGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGCTGCGCGGAGGCGCACTGC 1560
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1561 TGGGTCCAGGCGCCCACTGCTGTAACATGACCCAGTTCCTTCTGCGGCGCAGGAGTGC 1620
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 DB 1621 GTGGAGGAATGCCAGTACTGTCAGGGGTCTCCCGAGGAGTATGTGAATGCCAGGCAC 1680
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 DB 1681 TTGCGCTGCCACCTTGTGAGTGTGAGGTCAGGTCAGGAAATGCTCAGTACCTGTTTGG 1740
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600

Db 1741 GCTGACCAGTGTGGCCCTGTGCCCACTATTAAGACCCCTCCCTTCTCGGTGGCCCGCTGC 1800
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu 620
Db 1801 CCCAGCGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 1861 GGCCGATGCCAGCCTTGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1920
Qy 641 GlyCysProAlaGlnAtqAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 7
ABZ35744

ID ABZ35744 standard; DNA; 3768 BP.

XX ABZ35744;

07-FEB-2003 (first entry)

Human ERBB2 polynucleotide SEQ ID NO 52.

Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
protozoacide; gene expression; antisense; tumour; infection; Plasmodium;
virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
Hepatitis C virus; human papilloma virus; gene; ds.

OS Homo sapiens.

XX DE10100588-A1.

PN 18-JUL-2002.

XX 09-JAN-2001; 2001DE-1000588.

XX 09-JAN-2001; 2001DE-1000588.

XX (RIBO-) RIBOPHARMA AG.

XX Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX WPI; 2002-683450/74.

XX Inhibiting expression of target genes, useful e.g. for treating tumors,
by introducing into cells two double-stranded RNAs that are
complementary to the target -

XX Claim 13; Page 38-39; 100pp; German.

XX The invention relates to inhibiting expression of a target gene in a cell
by introducing at least two oligoribonucleotides (dsRNAi and II), both
with a double-stranded (ds) structure of at most 49 sequential nucleotide
pairs. At least part of one strand (S1, S2) of the ds structures in each
of dsRNAi and II are complementary to regions in the target gene. The
method uses antisense inhibition of gene expression using double stranded
RNA inhibition (RNAi). The method is particularly used to treat tumours
or infections, especially by Plasmodium or viruses/viroids (pathogenic on
humans, animals or plants). The method provides more effective inhibition
of expression than known methods using a single dsRNA, even at very low
concentrations. When dsRNA has at least one unpaired nucleotide at the
end, stability (and thus effective concentration in the cell) is
improved and efficiency can be increased further by pretreating the cells
with interferon. The present sequence is that of a target DNA of the
invention.

XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:

Pred. No.: 1.51e-229 Length: 3768

Score: 3628.00 Matches: 653

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
SEQ3 (1-653) x ABZ35744 (1-3768)
Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCCTCCTCGCCCTTGTGCCCCCGGAGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGACCCCAAGTGTGCACCGGCACACATGAAGCTGGGCTGCCCTGCCAGTCCCGAG 120
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnVal 60
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGGAGGAACCTG 180
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAACCTACCTACCTGCCCAATGCCAGCGCTGCTTCTCGCAGGATATCCAGAGGTG 240
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGTGCTCATCGCTCACACCAAGTGGAGGAGGTCCCACTGCAGAGCTCGG 300
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGGCAGGCCACCCAGCTCTTTGAGGACAATATGCCCTGGCCGTGCTAGACAAATGGA 360
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGCTGAACAATACCCCTGTGCACAGGGGSCCTCCCGCAGGAGGCTCGCGGAGCTG 420
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCACGCGAAGCCCGAG 480
Qy 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCTGCTACCAGGACACGATTTTGTGAAGAGACATCTTCCACAGAAACACCACTGGCT 540
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTTCCGATGTGAAG 600
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTTGTCAGAGCCTGACGCGCACTGTCTGT 660
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCCGGTGGCTGTGCCGCTGCAGAGGGCCACTGCCCACTACTGCTGCTCCATGAGCACTGT 720
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCCCTGCCCTCCACTTCAACAC 780
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCAACACACAGACAGACGTTTGGAG 840
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCCGAGGGCCCGGTATACATTGGGGCCCACTGTGTGACTGCTGCTCCC 900
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACCTACCTTTCTACGGAGCTGGGATCTCGCACCCCTCTGCTGCCCCCTGCACAACCAA 960
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACACGACGAGGATGGAAACACACGCGGTGTGAGAAGTGCAGCAAGCCTGTGCCCGA 1020


```
Db 241 CAGGGCTACGTGCTCATCGTCTACAACCAAGTAGGAGGTCTCCCACTGCAGAGCTGCGG 300
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnThrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTAGACAATGGA 360
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGTGAACATAACACCCCTGTACAGGGGCTCCCAAGAGGCTCGGGAGCTG 420
Qy 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGln 160
Db 421 CAGCTTGAAGCCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 480
Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCGTGTACCCAGGACAGATTTTGTGAAGAGACATCTTCCACAAGAACACCACTGGCT 540
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCGCTCTTCTCCGATGTGAAG 600
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTAGAGCTGTGAGCGCACCTGTCTGT 660
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCCGGTGGCTGTCGCCGTGCAAGGGCCACTGCCCACTGACTGCTGCCATGACAGTGT 720
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTCGCGCTGCACGGGCCCAAGCACCTCTGACTGCTGGCTGCCCTCCACTTCAACCCAC 780
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGTGTCACCTACAACACAGACACAGCTTTGAG 840
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCATATCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTGCC 900
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACACTACCTTTCTAGCGAGCTGGATCTCTGCACCTCGTCTGCCCTGCACAACCAA 960
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAGTGCAGAACGCCCTGTGCCCGA 1020
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1021 GTGTGCTATGCTTGGGCATGGACACTTCGAGAGGTGAGGCGAGTTACCACTGCCAAT 1080
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 ATCCAGGAGTTGTGTGCTGCAAGAAGATCTTTGGAGCCTTGGCATTTCTGCCGGAGAGC 1140
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1141 TTTGATGGGGACCCAGCCTCCAAACACTGCCCGCTCCAGCGAGAGCTCCCAAGTGT 1200
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1201 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGCGCGGACAGCCTGCCT 1260
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGACGAAATCTGCACAATGGCGCC 1320
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGCTGACCCCTGCAAGGGCTGGGCATCAGCTGGTGGGGCTGCCTCACTAGGGAA 1380
```

```
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGGCAGTGGACTGGCCCTCATCCACATACACACCACCACTCTCTCTGTGCACACGGTG 1440
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCCTGGGACCAAGCTCTTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
Qy 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGCAGCAGTGTGTGGCGAGGGCTGCGCTGCCACCAAGCTGTGCCGCCGAGGCACTGC 1560
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGGTCCAGGGCCCACTGTCAGTGTCAACTGCAAGCAGTTCCTTCGGGGCCAGAGTGC 1620
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCAAGGAGTATGTGAATGCCAGCACTGT 1680
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCCGTGCCACCTGAGTGTACGCCCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1740
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1741 GCTGACCAAGTGTGTGGCTGTGCCCACTATTAAGACCCCTCCCTTCTGCGTGGCCGCTGC 1800
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGTGTGAAACCTGACCTCTCTACATGCCATCCCACTCTGGAAGTTTCCAGATGAGGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GGCCCATGCCAGCCTTGGCCCACTCACTGCACCACTCTCTGTGTGTGGACCTGGATGACAAG 1920
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGAGCTCC 1959

RESULT 9
AADA3935
ID AAD43935 standard; cDNA; 3768 BP.
AC AAD43935;
XX
XX
DT 13-DEC-2002 (first entry)
XX
DE Human HER-2 cDNA.
XX
XX Transgenic animal; transgenic; mammary gland cell; HER2; tumour;
KW cancer; therapy; apoptosis; cytostatic; human; gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..3768
XX /*tag= a
XX /product= "Human HER2 protein"
XX
XX US2002035736-A1.
XX
XX 21-MAR-2002.
XX
XX 16-MAR-2001; 2001US-0811115.
XX
XX 16-MAR-2000; 2000US-189844P.
XX
XX (ERIC/) ERICKSON S.
XX (KING/) KING K.
XX (SCHW/) SCHWALL R.
XX
```

PI Erickson S, King K, Schwall R;
 XX
 DR WPI; 2002-401155/43.
 DR P-PSDB; AAE26349.
 XX
 PT New transgenic non-human mammal that produces detectable levels of a
 PT native human HER2 protein in its mammary gland cells, useful as tumor
 PT models for testing HER2-directed cancer therapies, and for identifying
 PT anticancer agents -
 XX
 PS Example 2; Page 24-26; 83pp; English.
 XX
 CC The invention relates to a transgenic non-human mammal that produces in
 CC its mammary gland cells detectable levels of a native human HER2 protein
 CC or its fragment. The transgenic animals are useful as tumour models for
 CC testing HER2-directed cancer therapies, and for identifying anticancer
 CC agents. The animals may also be used as source of cells which can be
 CC immortalised in culture, in screening for compounds that have potential
 CC as prophylactic or therapeutic treatments of diseases or disorders
 CC involving expression of HER2. The anti-cancer molecules are useful for
 CC inducing apoptosis or cell death of cancer cells. The present sequence
 CC is human HER-2 cDNA.
 XX
 SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
 Pred. No.: Length: 3768
 Score: 3628.00 Matches: 653
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

SEQ3 (1-653) x AAD43935 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 DB 1 ATGGAGCTGGCGGCTTGGCGCTGGGGGCTCCCTCCCTCGCCCTTTCGCCCCCGAGGCC 60

QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 61 GCGAGCACCACCAAGTGTGACCGGCACACATGAAAGTGGCGCTCCCTGCCAGTCCCGAG 120

QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 DB 121 ACCCACTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGAAACCTG 180

QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 DB 181 GAACCTCACCTACCTGCCCAACCAATGCCAGCCTGCTCCTTCTGCAGGATATCCAGAGGTG 240

QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 241 CAGGGCTACGTGCTCATCGCTCACAAACCAAGTGAAGCAGGTCCCACTGCAGAGGTGCGG 300

QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 301 ATTTGCGCAGGACCCACCTCTTTGAGGACAACTATGCTTGGCGTGTCTAGACAAATGGA 360

QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB 361 GACCCGCTGAACAATACCACCCCTGTACAGGGGCTCCCGAGGAGGCTCGGGAGGTG 420

QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 DB 421 CAGCTTCGAAGCCCTCACAGATCTTGAAGAGGGGTCTTGATCCAGCGAAACCCCGAG 480

QY 161 LeuCysTyrGlnAspThrIleLeuTriPlysAspIlePheHisLysAsnAsnGlnLeuAla 180
 DB 481 CTCTGCTACCAAGGACACGATTTTGTGGAGGACATCTTCCACAGAAACCAACCACTGGCT 540

QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 1621 GTGAGGAAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680

DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTCCACCCCTGTTCTCCGATGTGTAAAG 600

QY 201 GlySerArgCysTyrPgluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 601 GGCTCCCTCTGCTGGGAGAGAGTTCTGAGGATTTGTAGAGCTTACCGGCACTGCTGTGT 660

QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 DB 661 GCCGTGCTGTGCCCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720

QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 721 GCTCGCGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCCCTGCTCACTTCAACCA 780

QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 781 AGTGCATCTGTGAGCTGCACTGCCCACTGGTGCCTTACCAACACAGACACAGTTTGG 840

QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 841 TCCATGCCCAATCCCGAGGCGGTATACATTCGGCGCCAGCTGTGACTGCTGCTGCC 900

QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 901 TACAACTACCTTTCTACGGAGCTGGGATCTCTGACCTCTGCTGCCCTGCACAAACCA 960

QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 961 GAGGTGCACGAGAGGATGGAAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1020

QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1021 GTGTGCTATGCTGGGCATGGAGCACTTGCAGAGGTGAGGCGAGTTACCACTGCCAAT 1080

QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1081 ATCCAGGAGTTTGTGCTGCAAGAAAGATCTTTGGGAGCTGGCATTTTTCGCGGAGAGC 1140

QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB 1141 TTTGATGGGACCCAGCCTCCAACTGCCCGCTCCAGCCAGACAGCACTCCAAGTGT 1200

QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
 DB 1201 GAGACTCTGGAAGAGATCACAGTTTACCTATACATCTCAGCATGCCGCGACAGCTGCT 1260

QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1261 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTTCTGCACAATGGCGC 1320

QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTyrLeuGlyLeuArgSerLeuArgGlu 460
 DB 1321 TACTGCTGACCCCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGCTGCTGAGGGAA 1380

QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1381 CTGGCAGTGGAGCTGGCCCTCATCCACCAATACACCCCACTCTGCTGCTGCACACAGTG 1440

QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1441 CCCTGGGACCAAGCTCTTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCA 1500

QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1501 GAGGACGAGTGTGGGCGAGGGCTGGCCTGCCACCAAGCTGTGGCGCCGAGGCACTGTC 1560

QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1561 TGGGCTCCAGGGCCCACTGCTGCTCACTGCAGCCAGTTCCTTCGCGGCGCAGGAGTGC 1620

QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 DB 1621 GTGAGGAAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680

```

QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
DB 1681 TTGGCGTGCCACCTGAGTGTACAGCCAGAAATGGCTCAGTGACCTGTTTGGACCGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrIlyAspProPheCysValAlaAArgCys 600
DB 1741 GCTGACACAGTGTGGCGCTGTGGCCACTATAAGAGACCTCCCTTCTCGTGGCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu 620
DB 1801 CCCAGCGGTGAAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAsPlys 640
DB 1861 GGCGCATGCCAGCCTTGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
DB 1921 GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 10
AAD43986
ID AAD43986 standard; DNA; 3768 BP.
XX
AC AAD43986;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human Her2 antigen DNA.
XX
KW Human; immune response; T-helper cell epitope; chitosan; CTL response;
KW vaccine; prostate cancer; breast cancer; Her2 antigen; cytostatic;
KW immunostimulant; gene; ds.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 1..3768
FT /*tag= a
FT /product= "Human Her2 antigen"
FT sig_peptide
FT 1..69
FT /*tag= b
FT mat_peptide
FT 70..2250
FT /*tag= c
FT /product= "Mature human Her2 antigen"
XX
PN WO200234287-A2.
XX
XX
XX 02-MAY-2002.
XX
XX 26-OCT-2001; 2001WO-DK00705.
XX
XX 27-OCT-2000; 2000DK-0001606.
XX
XX 03-NOV-2000; 2000US-245166P.
XX
XX 18-JUN-2001; 2001DK-0000936.
XX
XX (PHAR-) PHARMEXA AS.
XX
XX
XX Beier AM, Gautam A, Mouritsen S;
XX
XX WPI; 2002-463339/49.
XX
XX P-PSDB; AAE26366.
XX
XX Inducing or enhancing an immune response against an antigen,
XX particularly cytotoxic T-lymphocyte responses, for treating or
XX ameliorating prostate or breast cancer, comprises administering the
XX antigen formulated with chitosan
XX
XX Claim 28; Page 85-90; 97pp; English.
XX
XX The invention relates to a method for inducing or enhancing an immune
XX response against a polypeptide antigen in an animal, including human.

```

CC The method comprises administering the polypeptide antigen or at least
 CC one variant which includes at least one first T-helper cell epitope that
 CC is foreign to the animal (foreign TH epitope) and is formulated with
 CC chitosan. The polypeptide antigen is weakly immunogenic or non-
 CC immunogenic. The invention is used as vaccine. The chitosan and
 CC polypeptide antigen or its variant are useful in the preparation of an
 CC immunogenic composition for inducing or enhancing an immune response,
 CC particularly CTL response, against the polypeptide or protein antigen.
 CC The method for inducing or enhancing an immune response is useful in
 CC treating or ameliorating cancer, e.g. prostate or breast cancer. The
 CC present sequence is human Her2 antigen DNA.

XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:

Pred. No.: 1.51e-229 Length: 3768
 Score: 3628.00 Matches: 653
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

SEQ3 (1-653) x AAD43986 (1-3768)

```

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
DB 1 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCCTCTCGCCCTTTGCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
DB 61 GCGAGCACCACAAGTGTGCACCGCACAGACATGAAGCTGGGCTGCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnVal 60
DB 121 ACCACCTGGACATGCTCCGCCACTCTACAGGGCTGCCAGGTGGTGCAGGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
DB 181 GAACACCTACCTGCTGCCACCAATGCCAGCTGTCTTCTTGCAGGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
DB 241 CAGGGCTACGTGCTCATCGCTCACACCAAGTAGGAGGTGCCACATGCAGAGCTCGCG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
DB 301 ATGTGGGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCCGTGTAGACAAATGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
DB 361 GACCCGCTGAACAATACCAACCCCTGTACAGGGGCTCCCGAGGAGCTTGGCGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
DB 421 CAGCTTGAAGCCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGSAACCCCGAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla 180
DB 481 CTCTGTACCAGGACACGATTTGTGAAGAGATCTTCCACAGAGACACACCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
DB 541 CTCACACTGATAGACACCAACCCGCTCTCGGGCTGCCACCCCTGTCTTCCGATGTGAAG 600
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
DB 601 GGCTCCCGCTGCTGGGAGAGAGTTCCTGAGGATTTGCAGAGCTTGACGCGCACTGTCTGT 660
QY 221 AlacGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
DB 661 GCCGGTGGCTGTGCCCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260

```

```
Db 721 |GCTGGCGGTGACAGGGCCCCAAGCACTCTGACTGCTGGCTGCTCCACTTCACAC 780
QY 261 |SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 |AGTGGATCTGTAGCTGACCTGCCAGCCCTGGTCACTACACACAGACAGTTTGAG 840
QY 281 |SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 |TCCATGCCCAATCCGAGGGCCGGTATACATTCCGGCCGAGCTGTGTGACTGCCGTGCC 900
QY 301 |TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 |TACAACTACCTTTCTACGGACCTGGGATCTTCACCTCTGCTGCCCTCGCACACCA 960
QY 321 |GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 |GAGGTGACAGCAGAGATGACACAGCGGTGTGAGAGTGCAGAGCCCTGTGCCCA 1020
QY 341 |ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValAlaValThrSerAlaAsn 360
Db 1021 |CTGTGCTATGCTGGCATGGCAGCATTGCGAGAGGTGAGGCACTTACCAGTGCCAAT 1080
QY 361 |IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 |ATCCAGGAGTTTGTCTGGCTGCAAGAGATCTTTGGGAGCTGGCATTTCTGCCGAGAGC 1140
QY 381 |PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400
Db 1141 |TTTGATGGGACCCAGCTTCCAACTGCCCCGCTCCAGCCAGCAGCAGCTCCAAGTGT 1200
QY 401 |GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1201 |GAGACTCTGGAAGATCACAGTTTACCTATACATCTCAGCATGCCGCGACAGCTGCCT 1260
QY 421 |AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 |GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTTCTGCACAAATGCGCC 1320
QY 441 |TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 |TACTCGTGACCTCGAAGGCTGGGCATCAGCTGGCTGGGCTGGCTCAGTCTGAGGNA 1380
QY 461 |LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 |CTGGCAGTGGACTGGCCCTATCCACCATACACCCACCTCTGCTTCTGTGCACAGGTG 1440
QY 481 |ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 |CCCTGGGACCACTCTTCGGAAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
QY 501 |GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 |GAGGACGAGTGTGGCGAGGGCTGGCTGCCACCACTGTGGCCCGCAGGCGCACTGC 1560
QY 521 |TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 |TGGGTTCAGGGCCCAACCACTGTGTCAACTCAGCCAGTTCTTCTCGGGCCAGGAGTGC 1620
QY 541 |ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 |GTGGAGGAATGCCGAGTACTCAGGGGCTCCCGAGGAGATGTGAATGCCAGGCACGTGT 1680
QY 561 |LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 |TTGCGCTGCCACCTTGAGTGTGAGCCCGACAGATGGCTAGTGACCTGTTTGGACCGGAG 1740
QY 581 |AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db 1741 |GCTGACCACTGTGTGGCTGTGCCCACTATAAGGACCCCTCCCTTCTGCGTGGCCCGCTGC 1800
QY 601 |ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGlu 620
|
```

```
Db 1801 |CCACGGGTGTGAACCTTGACCTCTCTATACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 |GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
|
Db 1861 |GGCGATGCCAGCCTTGCCCCATCAACTGCACCACTCTCTGTGTGGACCTGGATGACAAG 1920
QY 641 |GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
|
Db 1921 |GGCTGCCCGCCGAGCAGAGAGCAGCGCCCTCTGACGTCC 1959

RESULT 11
ABV78168
ID ABV78168 standard; DNA; 3768 BP.
XX
AC ABV78168;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human ERBB2 DNA SEQ ID NO 52.
XX
KW RNA inhibition; dsRNA1; gene expression inhibitor; oncogene; cytostatic;
KW virucide; protozoacide; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200255693-A2.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-EP00152.
XX
PR 09-JAN-2001; 2001DE-1000586.
PR 26-OCT-2001; 2001DE-1055280.
PR 29-NOV-2001; 2001DE-1058411.
PR 07-DEC-2001; 2001DE-1060151.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
DR WPI; 2002-590671/63.
XX
PT Inhibiting expression of target gene, useful e.g. for inhibiting
PT oncogenes, by administering double-stranded RNA complementary to the
PT target and having an overhang
XX
PS Claim 10; Page 142-143; 203pp; German.
XX
CC The invention relates to inhibiting expression of a target gene (I) in a
CC cell by introducing an inhibitory RNA (dsRNA1) having a double-stranded
CC structure of at most 49 consecutive bases. At least part of one strand
CC (as1) of dsRNA1 is complementary to (I) and at least one end of dsRNA1
CC has an overhang of 1-4 nucleotides. The method is used to inhibit the
CC expression of a wide range of genes, e.g. oncogenes, cytokine genes etc.
CC in humans, also genes in Plasmodium or in viruses or viroids that are
CC pathogenic for humans, animals or plants. Introducing an overhang into
CC dsRNA greatly increases effectiveness for inhibiting gene expression,
CC both in vivo and in vitro and also increases stability and thus the
CC effective concentration inside the cell. The present sequence is that of
CC a gene related to the invention.
XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 1,51e-229 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
```

SEQ3 (1-653) x ABV78168 (1-3768)

Db 1201 GAGACTCTGGAGAGATCACAGGTACCTATATACATCTCAGCATGCCGAGCAGCGCTGCCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTCAGCGTCTCCAGAACCTGCAAGTATCCGGGAGCAAAATCTGCACAAATGGCGCC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTCGCTCACTGAGGAA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGCAGTGGAGCTGGCCCTCATCCACCATACACCCACCTCTGCTGTCACACGGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCA 1500
QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGACGAGTGTGGGCGAGGCTGCGCTGCCACCAAGCTGCGCCCGAGGCGACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGCTCCAGGGCCCCACCACTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCGAGTACTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCGTGCACCCCTGAGTGTGAGCCCGAGCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db 1741 GCTGACCAAGTGTGGGCTGTGCCACATATAGGACCCCTCCCTTCTGCGTGGCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GGCGCATGCCAGCCTTGCCCATCAACTGCACCCACTCCTCTGTGGACTGGATGATGACAAG 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCCCGGAGCAGAGAGCCAGCCCTCTGAGCTCC 1959

RESULT 13
ABA92250
ID ABA92250 standard; cDNA; 3768 BP.
XX AC ABA92250;
XX AC ABA92250;
XX 17-JUN-2002 (first entry)
XX Human Her-2/neu cDNA.
XX Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
XX receptor; human; gene therapy; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..3768
XX /*tag= a
XX /product= "Her-2/neu"
XX WO200212341-A2.
XX PN 14-FEB-2002.
XX PD

XX 03-AUG-2001; 2001WO-US24283.
XX PF 03-AUG-2000; 2000US-0632507.
XX PR (CORI-) CORIXA CORP.
XX PA (SMIK) SMITHLINE BEECHAM BIOLOGICALS.
XX PI Cheever MA, Gheysen D;
XX WPI; 2002-241743/29.
XX DR P-PSDB; AAM51143.
XX Her-2/neu fusion protein for treating or preventing cancer by eliciting
XX or enhancing an immune response to the protein, has Her-2/neu
XX extracellular domain fused to Her-2/neu intracellular or
XX phosphorylation domain
XX Disclosure; Fig 15; 141pp; English.
XX The present sequence is that of human Her-2/neu oncogene cDNA.
XX The cDNA encodes Her-2/neu (p185), an oncogenic self protein and
XX target for anti-cancer vaccines. The Her-2/neu gene is amplified
XX and p185 is overexpressed in a variety of cancers, including breast,
XX ovarian, colon, lung and prostate cancer. Her-2/neu is a member
XX of the tyrosine kinase family of receptor-like glycoproteins.
XX Its overexpression correlates with a poor prognosis in breast and
XX ovarian cancers. The invention provides Her-2/neu fusion
XX proteins, nucleic acids encoding them, viral vectors, and vaccines
XX comprising the fusion proteins or nucleic acid molecules. In
XX preferred fusion proteins, the extracellular domain of a Her-2/neu
XX protein is fused to a Her-2/neu intracellular domain or
XX phosphorylation domain (or its Deltap fragment). An immune
XX response to Her-2/neu protein is elicited or enhanced by
XX administering the fusion protein in the form of a vaccine, or by
XX transfecting cells of an animal ex vivo with a nucleic acid
XX encoding the fusion protein, and delivering the transfected cells
XX to the animal. The fusion proteins, nucleic acids, and isolated
XX specific T-cells are useful for inhibiting the development of a
XX cancer, especially breast, ovarian, colon, lung or prostate cancer
XX in a patient. T cells that specifically react with a Her-2/neu
XX fusion protein can be used to remove tumour cells from a sample in
XX order to inhibit the development of cancer in a patient.
XX SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 1.51e-229 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
SEQ3 (1-653) x ABA92250 (1-3768)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProProGlyAla 20
Db 1 ATGGAGCTGGCGGCTGTGCGCTGGGGGCTCTCTCTCGCCCTCTGCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCGAGTCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCCACCTGGACATGCTCCCGACCTCTTACCAGGGCTGCCAGGTGGTGGGAAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 181 GAATCACCCTACCTGCGCCACCAATGCCAGCTGTCTCTCTGAGGATATCCAGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100

PA (CORI-) CORIXA CORP.
 XX Gaiger A, Cheever MA, Hand-zimmermann S;
 PI WPI; 2002-280741/32.
 DR P-PSDB; AAU77114.
 XX Inhibiting haematological malignancy development by administering
 PT polypeptide comprising immunogenic portion of Her-2/neu, polynucleotide
 PT encoding the polypeptide, or antigen presenting cells expressing the
 PT polypeptide -
 XX Disclosure; Page 66-71; 74pp; English.
 XX The invention relates to a method for inhibiting development of
 CC haematological malignancy in a patient by administering a polypeptide
 CC comprising an immunogenic portion of Her-2/neu or a polynucleotide
 CC encoding the polypeptide. Antigen presenting cells that express the
 CC protein can also be administered. The sequences are used for inhibiting
 CC development of haematological malignancy such as acute myelogenous
 CC leukaemia (AML), chronic myelogenous leukaemia (CML), chronic lymphocytic
 CC leukaemia (CLL), MDS, myelomas, Hodgkin's lymphoma and non-Hodgkin's
 CC lymphoma. This sequence represents DNA encoding human Her-2/neu
 CC polypeptide.
 XX
 SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.51e-229 Length: 3768
 Score: 3628.00 Matches: 653
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 SEQ3 (1-653) x ABK10730 (1-3768)
 QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProGlyAla 20
 DB 1 ATGGAGCTGGCGGCTTGTGCGGCTGGGGCTCTCTCGCCCTTTGCCGCCGGAGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 61 GCGAGCACCAAGTGTGCACCGGCACACATGAGCTGGGCTCCCTGCGAGTCCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 DB 121 ACCCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCAGGTGGTGCGAGGAAACCTG 180
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
 DB 181 GAACCTACCTACCTGCGCCCAATGCCAGCCTGTCTCTGCGAGGATATCCAGGAGTG 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 241 CAGGGCTACGCTCATCGCTACACCAAGTAGGAGGAGTCCCATGCGAGGCTCGG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 301 ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB 361 GACCCGCTGAACAATACACCCCTGTACAGGGGCTCCCGAGGAGGCTCGCGGAGCTG 420
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 DB 421 CAGCTTGAAGACCTCACAGAGATCTTGAAGAGGAGGGCTCTGATCCAGCGGAACCCCCAG 480
 QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAla 180
 DB 481 CTCTGCTACCCAGGACGATTTTGTGGAAGGACATCTTCCACAGAACACACAGCTGGCT 540

QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAG 600
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 601 GGCTCCGCTGCTGGGAGAGAGTCTGAGGATTTGTAGAGCCTGACGCGACTGCTCTGT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 DB 661 GCCGGTGGCTGCCCGCTGCAGGGGCCACTGCCACTGACTGCTGCCATGAGCAGTGT 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 721 GCTCCGGCTGCACGGGCCCAACGACTCTGACTGCTGGCTGCCCTCCACTTCAACCCAC 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGCTCACCTACACACAGACACGTTTGAG 840
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 841 TCCATGCCCAATCCCGAGGGCCGTATACATTCGCGGCCAGCTGTGTGACTGCTGTCCTC 900
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 901 TACAACCTACCTTTCTAGCGAGCTGGATCTCTGCCCTCGCTGCCCTGCACAAACCAA 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 961 GAGGTGACAGCAGAGGATGAACACACGCGGTGTGAGAAGTGCACGAAGCCCTGTGCCGA 1020
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1021 GTGTGCTATGCTGGCATGGAGCAGCTTCGAGAGGTGAGGGCAGTACCAGTGCACAT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1081 ATCCAGGAGTTGCTGGCTGCAGAGATCTTTGGAGCCTTGGCATTTCTGCCGAGAGC 1140
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB 1141 TTTGATGGGACCCAGCCTCCACACTGCCCGCTCCAGCAGAGCAGCTCCAAGTGTGT 1200
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpPrCAspSerLeuPro 420
 DB 1201 GAGACTCTCGAAGAGATCAGAGTTACTATACATCTCAGCATGCGCCGAGCAGCTGCCT 1260
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1261 GACCTCAGCCTCTTCCAGAACCTGCAAGTAAATCCGGGGGAGAAATCTGCACAAATGGG 1320
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCCCTCACTCAGGGAA 1380
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1381 CTGGCAGTGCAGTGGCCCTCATCCACCATTAACACCCACTGCTTCTCGTGCACACGGTG 1440
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1441 CCCTGGACCCAGCTCTTTCGGAACCCCGCACCAAGCTCTGTCTCCACACTGCCAACCGCCA 1500
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1501 GAGGACAGTGTGTGGGCGAGGGCTGGCTGCCACAGCTGTGCCCGCCGAGGACACTGC 1560
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1561 TGGGGTCCAGGGCCACCCAGTGTGTCACTGACGACGAGTCTCTTCGGGGCCAGGAGTGC 1620
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560

Db 1621 GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGAGTATGTGAATGCCAGGCACGTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCGGTGCCACCCCTGAGTGTGAGCCGAGATGGCTCAGTGACCTGTTTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1741 GCTGACCAGTGTGTGGCTGTGCCACTATAGGACCCCTCCCTTCTGGTGGCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleThrPheProAspGluGlu 620
Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GCGCGATGCCAGCCTTGCCCCATCACTGCACTCCCTCTGTGTGGACCTGGATGACAAG 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCGCGACAGAGACCGCCCTCTGACGTCC 1959

RESULT 15

ABL91709

ID ABL91709 standard; DNA; 3768 BP.

AC ABL91709;

XX 28-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 52.

XX Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW cytostatic; virucide; protozoacide; antibacterial; ds.

XX Homo sapiens.

XX DE10100586-C1.

XX 11-APR-2002.

XX 09-JAN-2001; 2001DE-1000586.

XX 09-JAN-2001; 2001DE-1000586.

XX (RIBO-) RIBOPHARMA AG.

XX Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX WPI; 2002-270454/32.

XX Inhibiting gene expression in cells, useful for e.g. treating tumors,
PT by introducing double-stranded complementary oligoRNA having unpaired
PT terminal bases

XX Claim 13; Page 41-42; 104pp; German.

XX The invention relates to a method for inhibiting expression of a target
CC gene (ABL91658-ABL91797) in a cell by introducing at least one
CC oligoribonucleotide that has a double-stranded structure consisting of at
CC most 49 sequential nucleotide pairs, with at least part of one strand
CC complementary with the target gene and has at least one end a
CC single-stranded segment of 1-4 nt. The method provides
CC oligoribonucleotides for antisense inhibition of gene expression useful
CC e.g. for treating tumours but the oligoribonucleotides may also be
CC directed against genes present in pathogens (e.g. Plasmodium or
CC viruses/viroids, pathogenic on humans, animals or plants) or against
CC cytokine, id, developmental or prion genes. The method provides more
CC effective inhibition of gene expression than use of known
CC oligonucleotides, probably because the unpaired overhang increases
CC stability and thus intracellular concentration.

XX SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
Alignment Scores:
Pred. No.: 1.51e-229 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ3 (1-653) x ABL91709 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGGAGCTGGGGCCCTTGTGCCCTGGGGGCTCTCTCGCCCTCTTGGCCCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCCCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu 60
Db 121 ACCCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAATCACCCTACCTGCCACCACCAATGCCAGCCTGTCTCTCCAGAGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGCTACGTGCTCATCGCTCACAAACCAAGTGAGCAGGTCCCCTGCAGAGGTGCGG 300
QY 101 lleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGCGAGGCGCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTAGACAATGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGCTGAACAATACCACCCCTGTACAGGGGGCTCCCCAGGAGGCTTGGGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTTCATCCAGCGGAACCCCG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCTGCTACAGGACACGATTTTGTGAAGGACATCTTCCACAAGAACAACACGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600
QY 201 GlySerArgCysTrpGlySerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTCACAGGCTTGCAGCGCCTGCTGCT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCCCGTGGCTGTGCCCGCTGCAAGGGGCACCTGCCCACTGCTGCTGCTGCTGCTGCTGCT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTCGCGGCTGCACGGGCCCCAAGCACTCTGACTGCTGGCTGCTGCTGCTGCTGCTGCT 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGAGCTGCATGCCCGCTGGTGCACCTACACACACAGACACGTTTGTAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCGAGGGCGGTATACATTGGCGCGCAGCTGTGTGACTGCTGCTGCC 900

Search completed: October 15, 2003, 23:49:41
Job time : 669.221 secs

Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACACCTTTCTACGGACGTGGGATCTGCACCCCTCGCTGCCCCCTGCACAACCAA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluIysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACAGCAGAGGATGGAACACACGGGTGTGAGAAAGTGCAGAAAGCCCTGTGCCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGlnValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGGTCTGGGCATGGAGCACTTCGGAGAGGTGAGGGCAGTTACCATGGCCAAT	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGGAGCCTGGCAITTTCTGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGATGGGACCCAGCCTCCAACTATGCCCGCTCCAGCCAGCAGCAGCTCCCAAGTGT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaIrrProAspSerLeuPro	420
Db	1201	GAGACTCTGGAGAGATCACAGGTTACTATACATCFACGATGGCCGAGACGCCCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGAGCGAATTCGTGCACATGGCGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTTGCAGAGGCTGGGCATCAGCTGGCTGGGCTGGCTCCTCCTGAGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGCGAGTGGACTGGCCCTCATCCACATTAACACCCACCTCTGCTGCTGTCACAGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCCAGCTCTTCGGAACCCGCCACCAAGCTCTGCTCCACACTGCCAACCCGGCCA	1500
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACCAAGTGTGTGGCGAGGGCTGGCGTGCCACCAGCTGTGCGCCGAGGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGGCCACCCAGTGTGTCACTCAGCAGCCAGTTCCTCTGGGGGCCAGAGTGC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGAAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCCTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCGGTGCCACCTTGAGTGTGACGCCCAAGATGGCTCAGTGACCTGTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys	600
Db	1741	GCTGACCAAGTGTGGCCTGTGCCCATATAAGAGACCCCTCCCTTCTCGCTGGGCCCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTGAACCTTGACCTCTCTACATGCCCATCTGGAAGTTCCTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCCAGCCTTGCCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGAGCAGAGCAGCCGCTCTCACGTCC	1959

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 9403.76 Seconds
(without alignments)
2840.777 Million cell updates/sec

Title: SEQ3
Perfect score: 3628
Sequence: 1 MELALCRWGLLLALLPPGA.....CVDLDDKGPQAEQASPLTS 653

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -P/HOLLERAN480/runat_15102003_131912_20536/app_query.fasta_1.4685
-Q/cgn2_1/USPTO.spool_P/HOLLERAN480/runat_15102003_131912_20536/app_query.fasta_1.4685
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pcpt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480_4CGN_1_1_22724_8runat_15102003_131912_20536 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3628	100.0	2385	6	AR082744 Sequence
2	3628	100.0	2385	6	AR099963 Sequence
3	3628	100.0	2385	6	AR143949 Sequence
4	3628	100.0	3768	6	AR034479 Sequence
5	3628	100.0	3768	6	AX060704 Sequence
6	3628	100.0	3768	6	AX201817 Sequence
7	3628	100.0	3768	6	AX380923 Sequence
8	3628	100.0	3768	6	AX384604 Sequence
9	3628	100.0	3768	6	AX465456 Sequence
10	3628	100.0	3768	6	AX467229 Sequence
11	3628	100.0	3768	6	AX481438 Sequence
12	3628	100.0	4473	6	AR080259 Sequence
13	3628	100.0	4473	6	AR167390 Sequence
14	3628	100.0	4473	9	HSEB2R
15	3628	100.0	4530	6	AR202597 Sequence
16	3628	100.0	4530	6	AR283481 Sequence
17	3628	100.0	4530	6	AX282577 Sequence
18	3628	100.0	4530	6	AX587649 Sequence
19	3628	100.0	4530	6	AX644071 Sequence
20	3628	100.0	4530	6	BD005474 Cellular
21	3628	100.0	4530	6	I21124 Sequence 9
22	3628	100.0	4530	6	I59745 Sequence 9
23	3628	100.0	4530	9	HUMHER2A
24	3628	100.0	9274	6	AX060703 Sequence
25	3525	97.2	2871	6	AX023363 Sequence
26	3471	95.7	1872	6	AR181000 Sequence
27	3458	95.3	3678	6	AX505114 Sequence
28	3324.5	91.5	3780	4	AB008451 Canis fam
29	3110.5	85.7	4727	10	AY116182
30	3106	85.6	3955	6	AX380924
31	3106	85.6	3955	6	I21129
32	3106	85.6	3955	6	I59750
33	3106	85.6	3955	10	RNNEUR
34	3104.5	85.6	2763	6	AX380942
35	3104.5	85.6	2781	6	AX380944
36	3100.5	85.5	3771	6	AX189662
37	3100.5	85.5	3771	6	AX380925
38	3100.5	85.5	4694	10	BC046811 Mus muscu
39	3100.5	85.5	4695	10	BC053078 Mus muscu
40	3095	85.3	4062	10	HAMNEU
41	1874.5	51.7	1316	9	AF177761
42	1706.5	47.0	157216	9	AC087491
43	1706.5	47.0	161815	9	AC079199
44	1706.5	47.0	168585	9	AC040933
45	1706.5	47.0	198008	2	AC142197

ALIGNMENTS

RESULT 1

AR082744
LOCUS AR082744 2385 bp DNA linear PAT 01-SEP-2000
DEFINITION Sequence 3 from patent US 5976546.
ACCESSION AR082744
VERSION AR082744.1 GI:10009534
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2385)
AUTHORS Laus, R., Ruegg, C., London, and Wu, H.
TITLE Immunostimulatory compositions
JOURNAL Patent: US 5976546-A 3 02-NOV-1999;
FEATURES
Location/Qualifiers
1..2385
/organism="unknown"
BASE COUNT 488 a 780 c 677 g 440 t
ORIGIN

Alignment Scores:
Pred. No.: 1..7e-226 Length: 2385
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

SEQ3 (1-653) x AR082744 (1-2385)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
DB 11 ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCTCTCGCCCTCTTGGCCCGGAGCC 70
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
DB 71 GCGAGCACCCAGTGTGCACCGCACGACATGAAGCTCGGCTCCCTGCCAGTCCCGAG 130
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
DB 131 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGAGGAAACCTG 190
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
DB 191 GAACCTACCTAGCTGCTCCACCAATGCCAGCTGTCTTCTCCAGGATATCCAGAGGTG 250
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
DB 251 CAGGGCTACGTGCTATCGCTCACACCAAGTGAAGGAGGCTCCACCTGCAGAGGTGCGG 310
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
DB 311 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGA 370
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyCysLeuArgGluLeu 140
DB 371 GACCCGCTGAACAATACCACTCTGTACAGGGGCTTCCCGAGAGGCTGCGGAGCTG 430
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
DB 431 CAGCTTCGAGGCTCACAGAGATCTTGAAGGAGGGGTCTTATCCAGCGGAACCCCGAG 490
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
DB 491 CTCTGCTACAGGACACGATTTTGTGAAGGACATCTTCCACAAGAACCAACAGCTGGCT 550
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
DB 551 CTCACACTATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAAG 610
QY 201 GlySerArgCysTrpGlyGluSerGluAspCysGlnSerLeuThrArgThrValCys 220
DB 611 GGCTCCCGCTGCTGGGGAGAGAGTCTTGAGGATTTGTACAGGCTGCAGGCTGCAGCGCTGTCTGT 670

QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
DB 671 GCCGGTGGCTGTGCCGCTGCAAGGGCCACTGCCACTGCTGCTGCCATGAGCAGTGT 730
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 731 GCTGCCGGCTGCACGGGCCCAAGCACTGCTGCTGCCTGGCTGCCTCCTCACTTCAACAC 790
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
DB 791 ACTGGCATCTGTGAGCTGCACGCCAGCCCTGGTACCTACACACACACACAGCTTTGAG 850
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
DB 851 TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGGCCAGCTGTGTGACTGCCGTGCC 910
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
DB 911 TACAACCTACCTTTCTACGGACGTGGATCCTGCACCTCTGCTGCCCTGCACACCAAA 970
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 971 GAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCCGA 1030
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1031 GTGTCTATGGTCTGGGCATGGACACTTGCAGAGGTGAGGCACTTACCAGTGCCTAAT 1090
QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
DB 1091 ATCCAGGATTTGCTGGCTGCAAGAGATCTTTGGAGGCTGGCATTTCTGCCGAGAGC 1150
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
DB 1151 TTGTATGGGACCCAGCTCCCAACACTGCCCGCTCCAGCCAGAGAGTCCAAGTGT 1210
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
DB 1211 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCCGACAGCTGCCT 1270
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
DB 1271 GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACAAATGGCGC 1330
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
DB 1331 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCCTCACTAGGGAA 1390
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
DB 1391 CTGGCAGTGGACTGGCCCTCATCCACCAATAACACCCACTCTGCTTGTGCACACGGTG 1450
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
DB 1451 CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACTGCCAACCGGCCA 1510
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
DB 1511 GAGGACGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGCTGGCCCGCGGACACTGC 1570
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
DB 1571 TGGGGTCCAGGGCCCAACCCAGTGTGTCAACTGCAGCAGTTCCTTCGGGGCCAGGAGTGC 1630
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
DB 1631 GTGGAGGAATCCGAGTACTGAGGGGCTCCCGCAGGGAGTATGTGAATCCAGGCACTGT 1690
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
DB 1691 TTGCCGTGCCACCCCTGAGTGTGACCCCCAGAAATGCTAGTACCTCTTTTGGACCGGAG 1750
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600

Db 1751 GCTGACAGTGTGGCTGTGCTCCCACTATAAGGACCTCCCTTCTGGTGGCCGCTGC 1810
 Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu 620
 Db 1811 CCCAGCGGTGGAACCTGACCTCTCCACATGCCCATCTGGAAGTTCCAGATGAGGAG 1870
 Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
 Db 1871 GCGCATGCCAGCTTGGCCCATCACTGACCCACTCTGTGTGGACTGGATGACAAG 1930
 Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 Db 1931 GGCTGCCCGCGGACGAGAGCCAGCCCTCTGACGTCC 1969

RESULT 2

LOCUS AR099963 2385 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 3 from patent US 6080409.
 ACCESSION AR099963
 VERSION AR099963.1 GI:12810411
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE
 AUTHORS Laus, R., Ruegg, C. Landon, and Wu, H.
 TITLE Immunostimulatory method
 JOURNAL Patent: US 6080409-A 3 27-JUN-2000;
 FEATURES Location/Qualifiers
 source 1..2385
 /organism="unknown"

BASE COUNT 488 a 780 c 677 g 440 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.7e-226 Length: 2385
 Score: 3628.00 Matches: 553
 Percent Similarity: 100.00 Conservative: 0
 Best Local Similarity: 100.00 Mismatches: 0
 Query Match: 100.00 Indels: 0
 DB: 6 Gaps: 0

SEQ3 (1-653) x AR099963 (1-2385)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
 Db 11 ATGGAGCTGGCGGCTTGTGGCGTGGGGGTCTCTCTCGCCCTTGGCCCGGAGGCC 70
 Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 71 GCGAGCACCCCAAGTGTGCACCGGCACACATGAGCTGGGCTCCCTGCCAGTCCCGAG 130
 Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValGlnGlyAsnLeu 60
 Db 131 ACCCACTGGACATCTCCGACCTCTTACCAGGGCTGCCAGTGGTGGCGGAAACCTG 190
 Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
 Db 191 GAATCACTACTCTCCCAACCAATGCCAGCTGTCTTCTCGAGGATATCCAGGAGGTG 250
 Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 251 CAGGCTAGCTGCTCATCGCTCACAAACCAAGTAGGCGAGGTCCCACTGCAGAGGCTGCGG 310
 Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 311 ATFTGCGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCGTGTAGACAATGGA 370
 Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 371 GACCCGTGAACAATACCACCCCTGTQACAGGGGCTCCCGAGGAGGCTCGGGAGCTG 430

Qy 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGln 160
 Db 431 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGAGGGGTCTTGATCCAGCGGAACCCCGC 490
 Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
 Db 491 CTCGTCTACCAAGGACACGATTTTGTGGAAGGACATCTTCCACAAAGAACACAGCTGGCT 550
 Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 551 CTCACATGATAGACACCAACCGCTCTCGGGCTCGCACCCCTGTCTCCGATGTGTAAAG 610
 Qy 201 GlySerArgCysTrpGlyLysSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 611 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGCAGAGCTTCAGCGCCTGCTCTGT 670
 Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 Db 671 GCCGTGGCTGTGCCCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 730
 Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 731 GCTGCCGCTGCACGGGCCCAAGACACTCTGACTGCTGGCTGGCTCCACTTCAACCCAC 790
 Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 791 AGTGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCCACTACAACACAGACACAGCTTGA 850
 Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 851 TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGACTGCTCTGCC 910
 Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 911 TACAACCTACCTTCTACGGAGCTGGGATCTTCACCTCGCTGGCTGGCTGGCTGGCTGG 970
 Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 971 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTGCGAGCAAGCCCTGTGCC 1030
 Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1031 GTGTGCTATGCTTGGGCATGGACACTTGGCAGAGGTGAGGGCAGTTACCAGTCCCAAT 1090
 Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 Db 1091 ATCCAGGAGTTTGTGGCTGCAAGAGATCTTTGGGAGCCCTGGCATTTCTGCCGGAGAGC 1150
 Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 Db 1151 TTTGATGGGACCCAGCCTCCCAACACTGCCCGCTCCAGCCAGCAGACAGCTCCCAAGTGT 1210
 Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 Db 1211 GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGCCCGGAGACGCTGCT 1270
 Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1271 GACCTCAGCTCTCCAGAACCTGCAAGTATATCCGGGAGCAGATTTCTGCACAAATGGCC 1330
 Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1331 TACTCGCTGACCTGCAAGGGCTGGGCATCAGTGGCTGGGCTCGCTCCTACTGAGGAA 1390
 Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1391 CTGGCAGTGACTGGCCCTCATCACCATTAACACCCACTCTGCTGCTGCTGCACACGCTG 1450
 Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1451 CCCGGGACCAAGCTCTTTCCGAACCCCGCACCAAGCTGTGCTCCACACTGCCCAACCGG 1510
 Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520

[illegible]

```
Db 1271 GACCTCAGCGTCTTCAGAACCTGCAAGTAATCGGGGACGAATTCGCACAATGGCGCC 1330
Qy 441 TyrSerLeuThrLeuGlnGlyLeuCluYleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1331 TACTCGCTGACCCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCTCACTGAGGGAA 1390
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1391 CTGGCAGTGACCTGGCCCTCATCCACCATACACCCACCTCTGCTTCGTCACACAGGTG 1450
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1451 CCCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCCAACGGGCA 1510
Qy 501 GluAspGluCysValGlyGluCluYleAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1511 GAGGACGAGTGTGGCGGAGGGCTGGCCCTGCCACCAAGCTGTGGCCCGGAGGCACTGC 1570
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1571 TGGGTCCAGGGCCCAACCACTGTCTCAACTGCAGCCAGTTCTTCGGGGGCCAGAGTGC 1630
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1631 GTGGAGGAATCCCGAGTACTCGAGGGCTCCCGCAGGAGTATGTGAATGCCAGGCACCTGT 1690
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1691 TTGGCGTCCACCCCTGAGTGTACGCCCGCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1750
Qy 581 AlaAspGlnCysValAlaCysAlaHisTrpLysAspProPheCysValAlaArgCys 600
Db 1751 GCTGACCACTGTGTGGCTGTGCCACTATAAGGACCCCTCCCTTCGGTGGCGCGCTGC 1810
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1811 CCCAGCGGTGTGAACCTTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 1870
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1871 GGCGCATGCCAGCCTTGCCCATCACTGACCCACCATCTCTGTGTGGACCTGGATGACAA 1930
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1931 GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1969
```

```
RESULT 4
AR034479 LOCUS AR034479 3768 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5869445.
ACCESSION AR034479
VERSION AR034479.1 GI:5950084
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 3768)
TITLES Cheever, M.A. and Disis, M.L.
JOURNAL Methods for eliciting or enhancing reactivity to HER-2/neu protein
FEATURES Patent: US 5869445-A 1 09-FEB-1999;
Location/Qualifiers
1. 3768
/organism="unknown"
BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN
```

Alignment Scores:
Pred. No.: 2.84e-226 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

```
DB: 6 Gaps: 0
SEQ3 (1-653) x AR034479 (1-3768)
Qy 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
Db 1 ATGAGAGCTGGGGGCTTGTGGCGCTGGGGGCTCTCTCTCGCCCTTTTCCCGCCCGAGGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGACCCCAAGTGTGCCCGGCACACATGAAGCTGCCGCTCCCTGCCAGTCCCGAG 120
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCCACCTGGACATGCTCCGCCACCTCTACACAGGCTGCCAGGTGGTGCAGGAAACCTG 180
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAACCTCACTACCTGCCCAACCAATGCCAGCTGTCTTCTTCCAGGATATCCAGAGGTG 240
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGTGCTCATCGCTCACACCAAGTGAAGCAGGTCCCACTGCAGAGGCTGGCG 300
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGGCAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGCTAGACAATGGA 360
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCGCTGGAACAATACCCCTGTACAGGGGCTCCCGCAGGAGGCTTCGGGAGGTG 420
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTCAAGACCTCACAGAGATCTTGAAGAGAGGGGTCTTGATCCACCGGAACCCCGAG 480
Qy 161 LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCTGCTACCAAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACCAACAGCTGGCT 540
Qy 181 LeuThrIleLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACATGATAGACACCAACCGCTCTCGGGGCTGCCACCCCTGTCTTCCCATGTGTAAAG 600
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTCTCGGGGAGAGAGTCTGAGGATTTGTACAGAGCTGCACCGCAGCTCTGT 660
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db 661 GCCGGTGGCTGTGCCCGCTGCAAGGGGCACTGCCCACTGACTGCTGCTCATGAGCAGTGT 720
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTCCGGCTGCACGGGCCCCCAAGCACCTCTGACTGCCCTGGCTGCCCTCCACTTCAACCCAC 780
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGTGAGCTGCATGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCGAGGGCGGTATACATTCGGGCGCAGCTGTGTGCTGCTGCTGCTGCC 900
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACCTACCTTCTACGGAGCTGGGATTCCTGCAACCCCTGCTGCTGCCCTGCCAACCA 960
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGCAGGAGGATGGANACACAGCGGTGTGAGAAAGTGCACAGACCTCTGTGCC 1020
Qy 341 ValCystyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
```

```
Db      1021  |||||GTGTGATGCTGGCGCATGGAGCACCTTGGAGAGGTGAGGCGAGTTACCACTGCCAAT 1080
Qy      361   ILeclncluphealaglyCysLysLysIlePheglySerLeuAlaPheLeuProGluSer 380
Db      1081  ATCCAGAGATTCTGGCTGCAAGAGATCTTTGGAGGCTGGCATTTCTGCCGGAGAGC 1140
Qy      381   PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
Db      1141  TTTGATGGGACCCAGCCTCCACACTGCCCGCTCAGCCAGAGAGCTCCAAGTGT 1200
Qy      401   GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db      1201  GAGACTCTGGAAGAGATCACAGGTTACTATACATCTCAGCATGGCGGACAGCCTGCCT 1260
Qy      421   AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db      1261  GACCTCAGCGCTCTCCAGAACCTGCAGTAATCCGGGGAGCAATCTGCACAAATGCGCC 1320
Qy      441   TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1321  TACTCGCTGACCTGCAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCCTCAGTGGGAA 1380
Qy      461   LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1381  CTGGGAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTG 1440
Qy      481   ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db      1441  CCTGGGACAGCTCTTTTCGGAACCGGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
Qy      501   GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db      1501  GAGACAGAGTGTGGCGGAGGCCCTGGCCCTGCCACACAGCTGTCGCCCGAGGCGACTGC 1560
Qy      521   TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db      1561  TGGGGTCCAGGGGCCACCCAGTGTCTCACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620
Qy      541   ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db      1621  GTGGAGAGTGTCCAGTACTGCAAGGGCTCCCGCAGGGAGTATGTGAATGCCAGGCACTGT 1680
Qy      561   LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db      1681  TTGCCGTGCCACCCTGAGTGTACGCCCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1740
Qy      581   AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db      1741  GCTGACCAAGTGTGGCCCTGTGCCACTATAAGGACCCCTCCTTCGCGTGGCCCGCTGC 1800
Qy      601   ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db      1801  CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTCGAAGTTCAGATGAGGAG 1860
Qy      621   GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db      1861  GGGCGCATGCCAGCTTGCCCATCACTGCACCCACTCCTCTGTGTGACCTGGATGACAAG 1920
Qy      641   GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db      1921  GGCTGCCCGCCGAGCAGAGAGCAGCCCTCTGACGTCC 1959

RESULT 5
AX060704
LOCUS      AX060704
DEFINITION Sequence 2 from Patent WO0100244.
VERSION    AX060704
KEYWORDS   AX060704.1 GI:12406101
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE 1
AUTHORS   Erickson,S. and Schwall,R.
TITLE      Methods of treatment using anti-erbB antibody-maytansinoid
            conjugates
JOURNAL    Patent: WO 0100244-A 2 04-JAN-2001;
            Genentech, Inc. (US)
FEATURES   Location/Qualifiers
            source          1..3768
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN
Alignment Scores:
Pred. No.:      2.84e-226      Length:      3768
Score:          3628.00      Matches:      653
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              6              Gaps:      0

SEQ3 (1-653) x AX060704 (1-3768)
Qy      1   MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db      1   ATGGAGCTGGGGCCTTGTGGCGCTGGGGCTCCTCTCGCCCTTGGCCCCGGAGCC 60
Qy      21   AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      61   GCGAGCACCAAGTGTGCACCGSCACAGACATGAAGCTCGGCTCCTGCCAGTCCCGAG 120
Qy      41   ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db      121  ACCCACTGGACATGTCTGCCACCTCTACCAAGGCTGCAGGTGTGCGAGGAAACCTG 180
Qy      61   GluLeuThrTyrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIle 80
Db      181  GAACCTCACCTACTTGCACCAATGCCAGCTGTCTCTCGAGAGATATCCAGAGAGTG 240
Qy      81   GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      241  CAGGGCTACGTGTCTATCGCTCACAAACCAAGTAGGACAGTCCCACTGCAGAGGCTGCGG 300
Qy      101  IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      301  ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCCTGTAGACAATGGA 360
Qy      121  AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      361  GACCCGCTGAACAATACCACTCTGTACAGGGGCTCCCGAGGAGCCTCGGGGAGCTG 420
Qy      141  GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db      421  CAGCTTCGAAGCTTCACAGAGATCTTGAAGAGAGGGTCTTGATCCAGCGGAACCCCCAG 480
Qy      161  LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      481  CTCTGCTACCAAGACACGATTTGTGGAAGGACATCTTCCACAAGAACCAACAGCTGGCT 540
Qy      181  LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      541  CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG 600
Qy      201  GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      601  GGCCTCCGCTGCTGGGGAGAGAGATTCTGAGGATTGTACAGAGCTGACGGCGCACTGTCTGT 660
Qy      221  AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db      661  GCGGTGGCTGTGCCCGCTGCAAGGGGGCCACTGCCCACTGCACCTGCCTGCTCCATGACAGCTGT 720
```

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1
Erickson,S. and Schwall,R.
Methods of treatment using anti-erbB antibody-maytansinoid
conjugates
Patent: WO 0100244-A 2 04-JAN-2001;
Genentech, Inc. (US)

Location/Qualifiers
source
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN

QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 721 GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCCTGGCTGCCCTCCACTTCAACCAC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
DB 781 AGTGGCATCTGTGAGCTGCACCTGCGCCAGCCCTGGTCACCTAACACACAGACAGCTTTGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
DB 841 TCCATGCCCAATCCCGAGGCGGTATACATTCGCGGCCAGCTGTGTGACTGCCTGCTCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
DB 901 TACAACATCTTCTACGACGCTGGGATCCTGCACCTCGTGTGCCCTCGCACACACAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 961 GAGGTGACAGAGAGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCCA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1021 GTGTGCTATGGTCTGGCATGGAGCACTTGCAGAGGTGAGGCGAGTTACCACTGCCAAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
DB 1081 ATCCAGAGCTTGTGCTGCTGCAAGAAGATCTTTGGGAGCTGGCATTTCTGCCGGAGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluLeuGlnValPhe 400
DB 1141 TTTGATGGGAGCCAGCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTTCAAGTGT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
DB 1201 GAGACTCTGGAAGATCACAGGTACCTATACATCTCAGCATGGCCGACAGCTGCTC 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
DB 1261 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGACGAATCTGCACAAATGGCGC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
DB 1321 TACTCGCTCAGCTTCCAGGGCTGGGCATCAGCTGGCTGGGCTGGGCTCAGCTGAGGGA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
DB 1381 CTGGCAGTGGACTGSCCTCATCCACCAATACACCCACCTCTGCTTCTGCACAGGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
DB 1441 CCCTGGGACCACTCTTCCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
DB 1501 GAGGACAGTGTGTGGCGAGGCGCTGGCTGCCACCACTGTGGCCCGGAGGCACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
DB 1561 TGGGCTCCAGGCGCCACCACTGTGTCACTGCAGCACTTCTTCTGGGCGCCAGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
DB 1621 GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTGTGTAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
DB 1681 TTGCGCTGCACCTCTGAGTGTGAGCCCAAGATGGCTCAGTGCCTGTTTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
DB 1741 GCTGACCACTGTGTGGCTGTGCCCACTATAGGACCTTCCCTTCTGCGTGGCGGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620

DB 1801 CCCAGCGGTGTAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
DB 1861 GCGCATGTCACGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
DB 1921 GCGTCCCGCGGAGCAGAGAGCCAGCCCTCTGACGCTCC 1959
RESULT 6
AX201817
LOCUS AX201817 3768 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0153463.
ACCESSION AX201817
VERSION AX201817.1 GI:15391666
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Cheever, M.A. and Hand-Zimmermann, S.
TITLE Compounds and methods for prevention and treatment of her-2/ neu associated malignancies
JOURNAL Patent: WO 0153463-A 1 26-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source 1..3768 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC60126.1"
/db_xref="GI:15391667"
/translation="MELAALCRWGLLLALLPPGAASQVCTGTMKRLPASPETHLD
MURHLVQGVQVQGNLELYLPTNASLFLQDIOEVQGVVLIAHNOVROVPLQRLRV
RGTOIFEDNYALAVLDNGDPLANTPTVTGASPGGLRELQRLRSITETLKGVLQRLNPQ
LCYQITLWKDIFHKNQALALHLDITNRSRACHPCSPCKGRCWGESSEDCQSLRT
VFCAGGCARKGPLPTDCCHEQCAAGCTGPKHSDCLACLHNHSGICELHCPALVTNT
DTFSMPNPEGRYTFGCACVTACTPYNLSDVGSCTLVCLHNEVTAEDGTQRCCK
SPKARVCYGLMEHLREVRVAVTSANIQFAGCKITFGSLAFILHNGAYSLTQGLGI
OPEQLQVFTLEETGYLYISAWPDLPSVFNQLOVIRGLTHNGAYSLTQGLGI
SWLGRSLRELGSGLALIHHTHLCFVHTVPMDLFRNPQALHTANRDEDCVCEG
LACHQLARGHCWGPGTQCVCNCSQFLRGQECVEECRVLQGLPREVYNARHCLPCHPE
CQPNGSVTCFGEADQCACAHYKDPFCVAKCPGVPKPLDLSMFIWPKPDEEGACQ
PCPINTHSCVDLDDKCPAEQASPLTSIIISAVVGLILVGLVGVVGLIILKRROKI
RKYTMRLIQETELVEPLTPSGAMPNQAMRIKETELRKVKVLGSGAGFTVYKGIWI
PDGENVKPIVAIKRENTSPKANKEILDEAYVMAGVGSPPYVSLGICLTSTVOLVT
OLMPYGCILLDHVRENRGLGSODLLNMCQIAKMSYLEDVRLVHRDLAARNVLKSP
NIHVKITDGLARLLDIDETEHADGKVPKIMWALESILRRRTHQSDVSWSGVTWNE
LMTFGAKPDYGIPIAREIPDLLEKBERLPOPICTIDVIMVMKMWIDSECRPRREL
VSEFSRMARDPQRFVYIQNEDLGPASPLSTFYRSLLDEDDMDGLVDABEYLVPOQGF
DLGMDAPAGAGGMVHHRRSSSTSGSGDLTLGLESEEEAPRSPPLAPSGAGSVDFDG
FCGMAAGKGLQSLPTHDRSPLOQYSEDPTVPLPSETDGYVAPLTCSPQPYVNPQDVR
POPSPRPGPLPAARPAGATLERPKTILSPGKNGVKDFAFGGAVENPEYLTPOGGA
PQHPFPFAPSPAFDNLIIYWDQPPRGPAPSTFKGTPTAENPEYLGLDVPV"
BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 2,84e-226 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ3 (1-653) x AX201817 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 Db 1 ATGGAGCTGGCGGCTTGTGGCGTGGGGCTCCTCCTCGCCCTTGTGCCCGGAGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 61 GCGAGCACCCCAAGTGTGACCGCGCACACATGATGAAGCTGCGGCTCCTGCGAGTCCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 121 ACCACCTGGAGATGCTCCGCCACTCTACACAGGCTGCCAGGTGGTGCAGGGAACCTG 180
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 181 GAATCACTACCTGCTCCGCCCAATGCCAGCGCTGCTTCCTGCAGGATATCCAGAGGTG 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 241 CAGGGCTACGTGCTCATGCTCACAACCAAGTGAAGGAGGCTTGTGATCCAGCGAACCAG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 301 ATTTGGCAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGCTGCTAGACAATGGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 361 CACCGCTGAACAATACCAACCCCTGTCAAGGGGCTTCCCGCAGAGGCTCGCGGAGCTG 420
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 Db 421 CAGTTTCAAGCCCTCAGAGATCTTGAAGAGGAGGCTTGTGATCCAGCGAACCAG 480
 QY 161 LeuCysTyrGlnAspThrIleLeuTriPlysAspIlePheHisLysAsnAsnGlnLeuAla 180
 Db 481 CTCGTACACAGGACAGATTTTGTGAAGACATCTTCCACAGAACCAACACCTGGCT 540
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 541 CTCACATGATAGACACCAACCGCTCTGGGCGCTGCCACCCCTGTTCTCCGATGTGAAG 600
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgValCys 220
 Db 601 GGCTCCCGCTCTGGGAGAGAGTCTGAGAGATTGTGAGCCCTGACCGCACCTGTCTGT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
 Db 661 GCCGTGGCTGTGCCCGCTGCAAGGGCCACTGCCACTGACTGCTGCCATGACAGTGT 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 721 GCTCCCGCTGCACGGGCCCCAAGCACTCTGACTGCTGCTGGCTGCCCTCCACTTCAACAC 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 781 AGTGCATCTGTGAGCTGCACTGCCAGCCCTGTGTCACCTTACACACAGACACCTTTGAG 840
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 841 TCCATGCCAATCCGAGGGCCGCTATACATTCGGCGCCAGCTGTGTGACTGCTGTGCC 900
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 901 TACAACATACCTTTCTACGGAGGTGGGATCTCTGACCCCTGCTGCTGCCCGCTGCACACCAA 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 961 GAGGTGACAGAGGATGGAAACACACGGGTGTGAGAGTGCACAGACCTGTGCCCGA 1020
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1021 GTGTGCTATGCTTGGCATGGACACTTGGAGAGGTGAGGCGAGTTACCAAGTGCCTAAT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380

Db 1081 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC 1140
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 Db 1141 TTTGATGGGACCCAGCCTCCAACTGCCCGCTCCAGCCAGAGAGCTCCAAGTGT 1200
 QY 401 GluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
 Db 1201 GAGACTCTGGAGAGATCACAGTTTACCTATACATCTCAGCATGGCCGAGAGCCTGCT 1260
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1261 GACCTCAGCTGTCTCCAGAACCTCAAGTAATCGGGGACGAATTTCTSCAAATGGCGCC 1320
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1321 TACTCGCTGACCTGCAAGGGCTGGCATCAGCTGGCTGGGGCTCGCTCACTCAGGGAA 1380
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1381 CTGGCAGTGGACTGGCCCTCATCCACCATTAACACCCACCTCTGCTTCTGCACACGGTG 1440
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1441 CCCTGGGACCAAGCTCTTTGGGAACCCGACCAAGCTCTGTCTCCACACTGCCAACCGGCCA 1500
 QY 501 GluAspGluCysValGlyGluLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 Db 1501 GAGACGAGTGTGTGGCGAGGGCTGGCCCTGCCACCAAGCTGTGCCCGCAGGACACTGC 1560
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 Db 1561 TGGGCTCCAGGGCCCAACCTGCTGTCACCTGACGACAGTTCCTTCGGGGCCAGAGTGC 1620
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 Db 1621 GTGAGGAATGCCAGTACTGCAGGGCTGCCAGGAGTATGTGAATGCCAGGACACTGT 1680
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1681 TTGCCGTGCCACCTGAGTGTGACGCCCCAGAAATGGCTCAGTGCAGTCTGTTTGGACCGGAG 1740
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 Db 1741 GCTCACCAGTGTGTGGCTGTGCCACTATATAAGACCTCTCCTTCTGCTGGCCCGCTGC 1800
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 Db 1861 GGCCGATGCCAGCTTGGCCCATCACTGACCCACTCTCTGTGTGGACCTGGATGACAAAG 1920
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 Db 1921 GGCTGCCCGCGAGCAGAGAGCCAGCCCTCTGAGCTCC 1959

RESULT 7

AX380923

LOCUS

DEFINITION

AX380923

ACCESSION

AX380923.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AX380923 3768 bp DNA linear PAT 18-MAR-2002
 Sequence 9 from Patent WO0212341.

AX380923

AX380923.1 GI:19575767

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Cheever, M. A. and Gheysen, D.

Her-2/neu fusion proteins

JOURNAL Patent: WO 0212341-A 9 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
Location/Qualifiers
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
CDS
/feature="human Her-2/neu protein"
/codon_start=1
/protein_id="CAD28392.1"
/db_xref="GI:19573768"
/translation="MELAAALRWGLLLALLPPGAASQTQVCTGTMKRLRLPASPETHLD
MRHLYQCVQVQVQVLELTYPTNASLFLQDIQEVQGVLIHAHQVQVQLQRLIRY
RGTQIFEDYLAVDNGDPLNNTPTVTGASPGRLQLRLSLTEILAGVYLIQRPQ
LCYQDTLWKDITFHNKQALTLIDNRSRACHPCSPMKGSRGSESDCSQSLTRT
VCGAGCARCKGLPPTDCHEQCAAGCTGPKHSACLALFHNHSGICEHLRCPALVTNT
DTFESMPNPEGRYTFGSCVTACPYNYLSTDVGSCTLVCPHLNQEVTAEDGTQCEKC
SKPCARVCYGLMEHLEVRNRAVTSANIQEFAGCKKIFGLAFLPESFDGDPASNTAPL
QPEQVVFETLEEITGYLYISAWPDSLPLSVFQNLQVIRGLRHNGAYSLSLQGLGI
SWIGLSRLGSLGIALIHNTHLGCFVHTVWDQLFRNPHQALLHANRDEDECYEG
LACHOLCARHGWCWGPQTCVNCQFQVLRGQCEVRCVLOGLPREYVNAHCLPCPE
COPQNGVTCGPEADOCVACAHYKDPFCVACRCPGVKPDLSYMPJWKFPDEEGACQ
PCPINCTHSCVDLDKGCPCAPQASPLISISAVVGLLVLVGVVFGVGLIKRRQKI
RKYTHRRQLQETELVEPLTPSGAMPNQAMRILKTELKRVKVLGSGAGFTYKGIWI
PDGENVKIPVAIKVLRNTSPKANKEILDEAYVMAGVSPYVSRLLGLTSLTVOLVT
OLMPTGCLLDHVRNRRGLGSODLLNMCQIAKMSYLEDVRLVHRDLAARNVLKSP
NHVKTDFGLARLLDIDETEVHADGKVPKIKMALESILRRFTHOSDVMSYGVTVWE
LMTFGAKPYDGIPIREIDLLEKGERLPQPPCTIDVYIMHWKMIIDSECRPREL
VSEFRMARDQRFVYIQNEDLGPASPLDSTFYRSLLDDDDMDDAEYLVPQGF
FCPDPAAGAGVHHRSSRSGGDLTLGLEPSEEEAPRSLAPSEAGSDVDFG
DLGMAAGLGLASLPHPSPQRYSEDTPVLPSETDGVAPLTCSPQPEVYNQPDFR
POPPSPREGPLPAAPAGATLERPKTLSPGKNGVVKDVFAGGAVENPEYLTPOGGA
POPHPPAFSPAFDNLVYWDQPPPERGAPPSTFKTPTPAENPEYLGLDVPV"
1..1959
misc_feature
/feature="extracellular domain (ECD) of human Her-2/neu"
2026..3765
misc_feature
/feature="intracellular domain (ICD) of human Her-2/neu"
2968..3765
misc_feature
/feature="phosphorylation domain (PD) of human Her-2/neu"
2968..3144
misc_feature
/feature="preferred portion of the phosphorylation domain
(delta PD) of human Her-2/neu"
759 a 1171 c 1119 g 719 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 2,84e-226 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ3 (1-653) x AX380923 (1-3768)
QY 1 MetGluLeuAlaLeuCysArgTrrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
DB 1 ATGGAGCTGGCGGCTTGTGGCGCTGGGGGCTCTCTCGCCCTCTTGGCCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
DB 61 GCGAGCACCCAAAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
DB 121 ACCCACTGGACATCTCCGACACCTTACCAGGCTGCCAGGTGGTGCAGGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
DB 181 GAACATCACTACTCTGCCACCACTGCCAGGCTGTCTTCTCCGACGATATCCAGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100

DB 241 CAGGCTACGTGCTCATCGCTCACCAACCAAGTAGGACAGTCCCACTGCAGAGGTGCGG 300
QY 101 lleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
DB 301 ATTGTGGAGGAGCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTAGACAAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
DB 361 GACCCGCTGAACAATACCAACCCCTCTCACAGGGGCTCCCCAGGAGGCTGCGGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
DB 421 CAGCTTCGAAGCCTCACAGATCTTGAAGAGAGGGGTCTTATCCAGCGGAACCCAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnGlnLeuAla 180
DB 481 CTCTGCTACAGGACACCATTTTGTGAAGGACATCTTCCACAAGAACACACCGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600
QY 201 GlySerArgCysTrrpGlySerSerGluAspCysGlnSerLeuThrArgThrValCys 220
DB 601 GGCTCCCGCTGCTGGGGAGAGAGTCTGAGGATTTCTCAGAGCTTGCAGCGCACTCTCTCT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
DB 661 GCGGCTGGCTGTGCCCGCTGCAAGGGGCACTGCCCACTGACTGTGCTGCAATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 721 GCTGCCGCTGCACGGGCCCCAAGCACCTCTGACTGCTGCTGGCTGCTTCCACTTCAACCC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
DB 781 AGTGCATCTGTGAGCTGCACTGCCAGCCCTGGTGGTCACTACACACAGACACAGTGTGAG 840
QY 281 SerMetProAsnProGluGlyArgTrrpPheGlyAlaSerCysValThrAlaCysPro 300
DB 841 TCCATGCCCAATCCCAGGCGGTATACATTCGGCGGCGAGTGTGTGACTGCCGTGCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
DB 901 TACAACTACCTTCTACGGAGCTGGGATCTTGCACCCCTCGTGTGCCCTTGCACAAACCA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTGCAGCAAGCCCTGTGCCCA 1020
QY 341 ValCysTyrGlyLeuGlyMetClnHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1021 GTGTCTATGGTCTGGGATGGAGCACCTTGCAGAGAGGTGAGGGCAGTTACCAAGTCCCAAT 1080
QY 361 lleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
DB 1081 ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCCGGAGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
DB 1141 TTTGATGGGACCCAGCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrrpProAspSerLeuPro 420
DB 1201 GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGCGCCGAGCAGCTGCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
DB 1261 GACCTCAGCTCTTCCACAACCTGCAAGTAAATCCGGGGAGCAAAATTTCTGCACAAATGGCC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460

```
Db 1321 TACTCGCTGACCCGCAAGGCTGGGCATCAGCTGGCTGGGCTGGCTGCTACTGAGGAA 1380
Qy 461 LeuGlySerGlyLeuAlaLeuLeuHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGCAGTGGAGCTGGCCCTCATCACCAATACACCCACCTGCTTCTGTCACACAGGTG 1440
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGACGAGTGTGGGGCAGGGCTGGCTGCCACCAAGCTGTGGCCCGGAGGCACTGC 1560
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGCTCAGGGCCACCCAGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluThrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCAGTACTCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT 1680
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGGCGTSCCACCCTGAGTGTACGCCAGAAATGGCTCAGTACCTGTTTGGACGGAG 1740
Qy 581 AlaAspGlnCysValAlaCysAlaHisTrpLysAspProPheCysValAlaArgCys 600
Db 1741 GCTACCAAGTGTGGGCTGTGGCCACTATAGGACCTCCCTTCTGCGTGGCCGCTGC 1800
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu 620
Db 1801 CCCAGCGTGTGAACCTGACCTCTCTACATGCCATCGCCATCTGGAAGTTTCCAGATGAGGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GGCAGTGCAGCTTGGCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAG 1920
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 8
AX384604
LOCUS AX384604
DEFINITION Sequence 1 from Patent WO0214503.
ACCESSION AX384604
VERSION AX384604.1 GI:19577806
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D., McNeill, P.D. and Vedvick, T.S.
TITLE Compositions and methods for the therapy and diagnosis of her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 1 21-FEB-2002; CORIXA CORPORATION (US)
FEATURES
source
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD28577.1"
/db_xref="GI:19577807"
/translation="MELAALCRWGLLIALLPPGAASVQVCTGTMKLRLLPASPETHLD MLRHLVGGCVQGNLELYLPTNGLSIPLFDIOEVGGVYLIAHNRQVPLORLRIV RGTQLFEDNVALAVLDNGDPLNNTTPVTGASPGGLRELQLRSLTEIKKGVLIRNPQ
```

```
LCYDTLLWKDIFHKNQALALTIDTNRSRACHPCSPMKSGRCWGESSEDCQSLIRT
VAGGARCKPLPTDCEHQAAGCTGPKHSDCLACLHFNHSGICELHRCALVITYNT
DTFSGMNPGRYTFGASCTACPYNTLSTDVGCTLVCPLHNOEVTAEADGTORCEK
SKPCARCYGIGMEHLREAVTSAHQEFAGCKIFGSLAFLESDEGDPASNTAPL
OPEQLQVETLEETGLYISANPDSLPDLISVFQNLQVIRIHLNGAYSLTQGLGI
SWLGRSLRELGSGLAIHNNTHLCFVHTVMDQLFRNPQHALLHTANRDECEGEG
LACHOLARGHCWGPCTQVCNCSQFLRGQECVECRVLQGLPREYNARHCLPCHPE
CQPNGSVTCFPGPADQCACAHYKDPFCVARCPGSKVLPYKFWFDEBAGACQ
PCNTCTHSCVDLDDKGCAPORASPLTSIIASVVGLLVVVLGVFGILLKRRQKI
RKYTMRLLOETELVEPLTPSGAMPNOAMEILKETELRKVKVLGSGAGFVYKGIWI
PDGENVKIPVAIKVLRNTPSPKANKEILDEAYVMAGVSPVSRLLGLTCLTSTVOLVT
OLMGLRSLREHNRNRLGSDLLNMCQIAKMSYLEDLVRLVHRLDAARNVLKSP
NHVITDFGLARLDDIDETEHADGKVPVKWMALESILRRRFTHQSDVWYGYVTWE
LMTFGARPDIGIPAREIPDLLEKGERLPQPICTIDVYIMVMKCMWIDSECRPREF
VSEFSRARDPQRFVIONDELGPASPLDSTFYSLDDDDMDVLAEEVLYVQQGF
FCDDPAPGAGMVHHRSSSTRSGGDLJLGLPSEEPAPRSLPASEGAGSYFDG
DLGMAAKGLQSLPTHPDSPLORYSEDTVPPLPSETGYVAFPLFCSPQPEIYNPDVR
POPPSREGPLPAARPAATLERKTLSPGKNVYKDFVAFPGGAVENPEYLTPOGGAA
POPHPPAFSPAFDNLVYWDQDPERGAPSTFKGTPTAENPEYLGIDVPV"
```

BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN

Alignment Scores:
Pred. No.: 2.84e-226 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

SEQ3 (1-653) x AX384604 (1-3768)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGGAGCTGGGGCTTGTCCGCTGGGGCTCTCTCGCCCTTGTCCCGCGGAGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGACCCCAAGTGTGCACGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCGAG 120
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGGAGGAACCTG 180
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAATCCTACCTACCTGCCACCAATGCCAGCTGCTCTCTCGAGGATATCCAGAGGTG 240
Qy 81 GlnGlyTyrValLeuLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGTGCTCATCGCTCACCAACCAAGTGCAGGCTGCCACTGCCAGCTGCGG 300
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGGAGGACCCAGCTCTTTGAGGACAATATGCTCCCTGGCCGTGTAGACAATGGA 360
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGCTGAAACATACCCCTGTCTACAGGGGCTCCCGAGGAGGCTGCGGGAGCTG 420
Qy 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyGlyValLeuLeuGlnArgAsnProGln 160
Db 421 CAGCTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGCAACCCAG 480
Qy 161 LeuCysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnGlnLeuAla 180
Db 481 CTCTGCTTACGAGCACAGATTTTGTGGAAGGACATCTTCCACAGAACAACACAGCTGGCT 540
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATACACACCAACCCCTCTCGGGGCTTCCACCCCTGTTCTCCGATGTGAAG 600
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220

Db	601	GCTCCCGCTGCTGGGGAGAGATTCTGAGGATTGTCAGAGCCTGACGCGCACTGCTGT	660
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	661	GCCGGTGGCTGTCGCCGCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT	720
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	721	GCTGCCGGCTGCACGGGCCCAAGCACCTCTGACTGCCTGGCGCTGCTCCACTTCAACCAC	780
QY	261	SerGlyTleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	781	AGTGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACCTACAACACAGACACGTTGAG	840
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCCGAGGGCCGGTATACATTGGCGCCAGCTGTGTGACTGCCCTGTC	900
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACTACCTTTCTACGGACCTGGGATCTGCACCTCGTCTGCCCTCGCACACCA	960
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1020
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGGTCTGGGATGGAGCATTGCCAGAGGTGAGGCGATTACCACTGCCAAT	1080
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGGAGTTTGTGCTGGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCCGAGAC	1140
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGTATGGGACCCAGCCTCCAACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1200
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCCGACAGCTGCT	1260
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCTCTCCAGAACCTGCAAGTAATCCGGGAGCAATTCGCACAATGGCGC	1320
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCTCAGTGAGGAA	1380
QY	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGCAGTGGACTGGCCCTCATCCACCATACACCCACCTCTGCTCGTCACACGGTG	1440
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCACTCTTTCCGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCA	1500
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGCTGGCGCCGAGGCACTGC	1560
QY	521	TrpGlyProGlyProThrGluCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGTCCAGGGCCCAACCACTGCTCACTGACCACTGCTTCTCGGGGCCAGAGATGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATCCGAGTACTGACGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCGCTGCCACCTCGAGTGTGAGCCGCCAGAAATGGCTAGTGACCTGTTTGGACCGGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCACTGTGTGCGCTGTGCCCACTATAGGACCTCCCTTCTGCGTGGCCCGCTGC	1800
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTAACACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Db	1861	GCGCATGCCAGCTTGGCCCATCACTGACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GCTGCCCGCCGACAGAGACCGCCCTCTGACGTCC	1959
RESULT 9			
AX465456			
LOCUS			
Sequence 1 from Patent WO0213847.			
AX465456			
VERSION			
AX465456.1 GI:21899819			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1			
Gaiger, A., Cheever, M.A. and Hand-Zimmermann, S.			
Methods for diagnosis and therapy of hematological and			
virus-associated malignancies			
Patent: WO 0213847-A.1 21-FEB-2002;			
CORIXA CORPORATION (US)			
FEATURES			
source			
1..3768			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
1..3768			
/note="unnamed protein product"			
/codon_start=1			
/protein_id="CAD42363.1"			
/db_xref="GI:21899820"			
/translation="MELALCRWGLLLALLPPCAASTQVCTGDMKRLRLPASPEHLD			
MLRHLQGCQVVGQNLITYLPTNASLSFLQDIQEVQGYVLTAHNOVROVPLQRLRV			
RGTLQFDNYALAVLDNDPLNNTPTVTGASPGGLREQLRLSLTILKGGVLIQNRPO			
LCYQDTILKDFHKNQLALTLDITNRSRACHPCSPMKSGRSGESSEDCQSLTRT			
VCAGCARCKGKPLPTDCCHEOCAGCTGPKHSDCLACLFHNSGICELHCPALVYNT			
DTFESMPNPGRYTFGASCTVACPNYLSTDVGSCTLVCLPHNOEVTADGTCQCEK			
KPCARVCYGLGMEHLREVRAVTSANIQEFAGCKKIFGSLAPLPSFGDDPASNTAPL			
QPEQLQVETLEEITGYLITISAMPDLPDLSVFQNLQVIRGRLLHNGAYSLTQGLGI			
SWLGLRSLEILGSLALITHNTHLCFVHTVPMDFRNPQHALLTANRDEDECYEG			
LACHQLCARGHCWGPGTCVNCQSQFLRGQCEBECRVLQGLPREYVNAHRLCPGCE			
CQPNQSVTFCFGEADQCAHAHYKPPFCVACRCPGKVPDLSYMPIWKFPDDEGAC			
PCPLNCTHSCVDLDDKGPAGEORASPLTSIIISAVVILLVVLVGVVFGILLKRRQKI			
RKYMRRLLQETELVEPLTPSGAMPNQAMRLKETELRKVKYVLGSGAGFYVKGWI			
PDGENKIPVAKLVURENTSPANKELIDEATVMAGVSPVVSRLUGLCILSTVOLVT			
QLMPYQGLDHVRENRRIGSQDLLNWCMTAKMSYLEDVRLVLRHDLAARNLVKSP			
NHVKTTFGLARLLIDETEHADGKVP IKWMALESILRRRFTHQSDVSGYVTWE			
LMTGAKPYDGIIPAREIPDLLEKGRLPQPTCTIDVYIMVVKMIDSECRPPREL			
VSESRMARDPQRFV IONEDLGPASPLDSTFYRSLLEDDMDGLVDAEYLVQOQF			
FCPDPAFGAGMVRHRRSSSTRSGGDLTLGLEPSEEA PRSPLAPSEGAGSDVFDG			
DLGMAAGLQSLPDPSPLOYSSEDPVLPKPSSETDGYVAPLCTSPQPEYVNPQDVR			
PQPPSRPGPLPAAPAGATLERPKTLPSPKGVVVKDFAFGAGVENEPEYLTPOGAA			
PQPPPPAFSPAFNLXYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVPV"			
BASE COUNT			
ORIGIN			
759 a 1171 c 1119 g 719 t			
Alignment Scores:			
Pred. No.:			
Score:			
2.84e-226 Length: 3768			
3628.00 Matches: 653			

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 SEQ3 (1-653) x AX465456 (1-3768)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
 Db 1 ATGGAGCTGGGGGCTTGGCGCTGGGGCTCCTCGCCCTCTTGCCTCGGAGCC 60

Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 61 CGAGACCCCAAGGTGACCGGCACAGATGAAGTGGCGCTCCCTGCGAGTCCCGAG 120

Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu 60
 Db 121 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGTGGTGCGAGGAACCTG 180

Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 181 GAACCTACCTACCTGCCACCAATGCCAGCTGTCTCTCTGCGAGATATCCAGGAGGTG 240

Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 241 CAGGGCTACGTGCTCATCGCTCAACCAAGTAGGAGGTGCCACTGCAGAGGCTGGG 300

Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 301 ATTGTGCGAGCACCCAGCTCTTTGAGGACAACATATGCCCTGGCGTGTCTAGACAATGA 360

Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 361 GACCGCTGAACAATACCAACCCCTGTACAGGGGCTCCCGAGGAGGCTCGCGGAGCTG 420

Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 Db 421 CAGCTTGCAGCCCTCAGAGATCTTGAAGAGGAGGCTTGTATCCAGCGGAACCCCGAG 480

Qy 161 LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 Db 481 CTCTGTACACAGGACAGATTTTGGNAGGACATCTCCACAGACCAACACCTGGCT 540

Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 541 CTCACCTGTATGATACCAACCCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600

Qy 201 GlySerArgCysTrpGlyGluSerSerClnuAspCysGlnSerLeuThrArgThrValCys 220
 Db 601 GGCTCCCGCTCTGGGAGAGAGTCTGAGGATTGTGAGAGCCTGACGCGCACCTGTCTGT 660

Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 Db 661 GCCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCACTGACTGCTGCGCATGACAGTGT 720

Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 721 GCTCGGGCTGCACGGGGCCCCAAGCACTCTGACTGCGCTGGCGTGCCTCCACTTCAACAC 780

Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 781 AGTGGCATCTGTAGCTGCACTGCCCGCCCTGTGTCACCTTACACACAGACACACTTTGAG 840

Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 841 TCCATGCCCAATCCGAGGGCCGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTGCC 900

Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 901 TACAACCTACCTTTCTACGAGCTGGGATCTCTGCAACCTCGCTCGCTGCCCAACCAA 960

Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340

Db 961 GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAAGTGACAGCAAGCCCTGTGCCCGA 1020

Qy 341 ValCysTyrGlyLeuGlyMetGlnHisLeuArgGluValArgAlaValThrSerAlaAsn 360

Db 1021 GTGTGCTATGCTTGGGCATGGAGCACTTGGCAGAGGTGAGGGCAGTTACCAAGTCCCAAT 1080

Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380

Db 1081 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTTGCCTGCGCAGAGC 1140

Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400

Db 1141 TTTGATGGGACCCAGCCTCCAACTGCCCGCTCCAGCCAGAGCACTCCAAGTGT 1200

Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420

Db 1201 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGCCCGGACAGCCTGGCT 1260

Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440

Db 1261 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGACGAATTTGCACAATGGCGCC 1320

Qy 441 TyrSerLeuThrLeuGlnGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460

Db 1321 TACTCGCTGACCCGCAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCTCTACTGAGGGAA 1380

Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480

Db 1381 CTGGCAGTGGACTGGCCCTCATCCACATTAACACCCACTCTGCTGTGTGCACACGGTG 1440

Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500

Db 1441 CCTTGGGACCAAGCTCTTTCGGAAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500

Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520

Db 1501 GAGACAGAGTGTGGGCGAGGGCTGGCCCTGCCACCAAGCTGTGGCGCCGAGGCACTGC 1560

Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540

Db 1561 TGGGGTCCAGGGCCCAACCCAGTGTCAACTGCAGCAGCACTTCTTCGGGGCCAGGAGTGC 1620

Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560

Db 1621 GTGGAGGAATCCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680

Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580

Db 1681 TTGCCGTGCCACCCGTGAGTGTGACGCCCGAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1740

Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600

Db 1741 GCTACACAGTGTGGCTGTGGCCACTATTAAGACCCCTCCCTTCTGCTGGCGCCGCTGC 1800

Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620

Db 1801 CCCAGCGTGTGAACCTGACCTCTCTCTACATGCCCATCTGGAAGTGTCCAGATGAGGAG 1860

Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640

Db 1861 GGCCGATGCCAGCCTGCCCCATCACTGCACCCACTCTCTGTGTGGAGCTGGATGACAAG 1920

Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653

Db 1921 GGCTGCCCGCCGAGCAGAGAGCCGCTCTGAGCTCC 1959

RESULT 10
 AX467229
 LOCUS
 DEFINITION Sequence 3 from Patent WO0234287.
 ACCESSION AX467229
 VERSION AX467229.1 GI:21900511
 KEYWORDS

AX467229 3768 bp DNA 1 linear PAT 16-JUL-2002

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Beler, A.M., Gautam, A. and Mouritsen, S.R.
TITLE       Novel therapeutic vaccine formulations
JOURNAL     Patent: WO 0234287-A 3 02-MAY-2002;
            Pharmexa A/S (DK)
FEATURES
  Source     1...3768
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
  CDS        1...3768
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD42452.1"
            /db_xref="GI:21900512"
            /translation="MELAAICRMGLLLALLPPGAASQTQVCTGTDKMLRLPASPEHLD
            MRHLQGVOCVGNLELTVPNTTPTVTSGLSPGGLRELQRLSLTEILKGGVLIQRPQ
            RGTOLFEDNVALVDNGDPLNNTPTVTSGLSPGGLRELQRLSLTEILKGGVLIQRPQ
            LCYQDTILKWDIFHKNNQALTLIDTNRSRACHPCSPWKGSRGSGESDCQSLTRT
            VCAGGCARCGPLPTDCCHEOCAGCTGPKHSDCLACLFHNSGICELHCPALVYNT
            DTFSMPNPGRYTFGASCTFACPNYLSTDVGSCTLVCPLNHOEVAEDGTQRCCK
            SKPCRCVYIGIMEHLREAVTSANIQEFAGCKIFGSLAFLPESDGDPAASNTAPL
            QPEOLVFETLEETGYLITISAWPDSLPLDSVFONLOVIRGRILHNGAYSILTLOGLI
            SWGLSLRELGSGLIHNTHLCFVHTVPMQDLFRNPHQALLHTANRDEDCVGBG
            LACHQLARGHCWGPQTQVNCQSFRLGQECVEECRVLQGLPREYVNRHCLCPHE
            CQPQNSVTCFPGPADQCVACAHYKDPFCVARGPSGVKPLDSYMPLWKPFDEGAQO
            PCPLNTHSCVDLDDKCPAEPASPLTSIVSAVGLLVGLVGVFGILTKRQQR
            RKTMYRLLEDELVEPLTPSPGPNQMRILKETELRKVKVLGSGAFGVYKGIWI
            PDGENVPIVAIKLRENTSPKANKEILDEAYVWAGVSPVSRLLGLTSTVOLVT
            QLMYPGCLDVRNENRGLSQDLLNMCQIAKMGVLEQVLRVHRLAARNVLKSP
            NHVKTDFGLARLDDIETEHADGKVPFKWMALESILRRRFTHQSDVNSGYTVWE
            LWFEGARPDGPAREIPDLLEGERLPQPPCTIDVIMVKNWCMIDSECRPFREL
            VSEFRWARDQPREVIONEDLGPASPLDTSFYRSLLDDDDMDGLVDAEYLYPQGF
            FCDPAPAGAGMWHRRSSTRSGGDLTLGLEPSEEPAPRSLAPSEAGSVDVFG
            DLGMGAAGLQSLPTDPSPLQYSEDTPVLPSETDGYVAPLTCSPQPEVNVQPDYR
            POPSPREGPLPAARPAQATLERAKTLSPGKNGVVKDVFAGGAVENPEYLTPOGGAA
            POPHPPAFSPAFDNLVYWDQDPPERCAPSTFKGTPTAENPEYLGLDVVP"
BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN
Alignment Scores:
Pred. No.:      2,84e-226      Length:      3768
Score:          3628.00        Matches:      653
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:     100.00%      Indels:        0
DB:              6            Gaps:         0

SEQ3 (1-653) x AX467229 (1-3768)
Qy      1 MetGluLeuAlaLeuCysArgTrrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
Db      1 ATGGAGCTGGGGCTTGTGCGCTGGGGCTCTCTCGCCCTCTTGGCCCGGAGCC 60
Qy      21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      61 GCGAGACCCCAAGTGTGCACCGGCACACATGAAGCTGGGGCTCCCTGCCAGTCCCGAG 120
Qy      41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db      121 ACCACCTGGACATGCTCCGCCACCTCTACCGAGGCTGCCAGGTGGTGACGGGAACCTG 180
Qy      61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnGlnGluVal 80
Db      181 GAACCTACCTACCTGCCACCAATGCCAGCTGCTCTTCTTCTGCAGGATATCCAGGAGGTG 240
Qy      81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      241 CAGGGCTACGTGCTCATCGCTCACACCAAGTGGAGGAGTCCCACTGCAGAGCTGCGG 300

```


QY	461	LeuGlySerGlyLeuAlaLeuLeuHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGCAGTGGAGTGGCCCTCATCCACCAATAACACCACCTCTGCTTCGTGCGACACGGTG	1440
QY	481	ProTirAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCACTCTTTCGGAACCCGCCACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500
QY	501	GlusAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACGAGTGTGGCGAGGCGCTGGCCCTGCCACAGCTGTGGCGCCGAGGCACTGC	1560
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGCCCAACCCAGTGTGTCAACTGCAGCCAGTCTCTCGGGGCCAGGAGTGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCGGTGCCACCTGAGTGTGACGCCAGAAATGGCTCAGTACCTGTTTGGACCGGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys	600
Db	1741	GCTGACCAAGTGTGGCCTGTGCCACTATAAGACACCTCCCTTCTCGTGGCGCGTGC	1800
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu	620
Db	1801	CCGAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGGCGATGCCAGCTTGCCTCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGTGTCCCGCCGACGACAGAGCCAGCCCTCTGACGTCC	1959
RESULT.11			
LOCUS	AX481438	3768 bp	DNA
DEFINITION	Sequence 52 from Patent WO02055693.		
ACCESSION	AX481438		
VERSION	AX481438.1	GI:22316352	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.		
TITLE	Method for inhibiting the expression of a target gene		
JOURNAL	Patent: WO 02055693-A 52 18-JUL-2002;		
FEATURES	Ribopharma AG (DE)		
source	Location/Qualifiers		
	1..3768		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	758 a	1170 c	1121 g 719 t
ORIGIN			
Alignment Scores:			
Pred. No.:	2.84e-226	Length:	3768
Score:	3628.00	Matches:	653
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
SEQ3 (1-653) x AX481438 (1-3768)			

Qy	1	MetGluLeuAlaLeuCysArgTrrpGlyLeuLeuAlaLeuLeuProProGlyAla	20
Db	1	ATGGAGCTGGCGGCTTGTCCGCTGGGGCTCTCTCTCGCTTCTTTCGCCCGGAGCC	60
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	61	GCAGACACCAAGTGTGCACCGGCACAGATGAAGCTGGGCTCCCTGCGAGTCCCGAG	120
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu	60
Db	121	ACCCACTGGACATGCTCCGCCACTCTACAGGGCTGCCAGGTGTGTGCGAGGAACCTG	180
Qy	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	181	GAACCTACCTACCTGCCCAATGCCAGCTGTCTCTCTCCAGGATATCCAGAGGTG	240
Qy	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	241	CAGGGCTACGTGCTCATCGCTCACAAAGTGAAGGAGGTCCCACTTCAGAGGCTCGCG	300
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	301	ATTGTGGGAGGACCCAGCTCTTTGAGGACAATATGCCCTGGCCGTGCTAGACAAATGA	360
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	361	GACCCGTGAACAATACCCCTGTGCACAGGGGCTCCCAAGAGGAGGCTCGCGAGCTG	420
Qy	141	GlnLeuArgSerLeuThrGluLeuLeuGlyGlyValLeuIleGlnArgAsnProGln	160
Db	421	CAGCTTCAAGACCTCACAGAGATCTTGAAGAGGGGTCTTGATCCACGCGAAGCCCGAG	480
Qy	161	LeuCysTyrGlnAspThrIleLeuTrrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	481	CTCTGTACACGAGACAGATTTGTGAAGAGCATCTTCACAGAACACACAGCTGGCT	540
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	541	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAAG	600
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	601	GGTCCCGCTGTCTGGGAGAGAGTCTGAGGATTTGTACAGACCTGAGCGCACTGTCTGT	660
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys	240
Db	661	GCCGGTGGCTGTGCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCTGATGACAGCTGT	720
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	721	GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT	780
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATCTGTGAGCTGCTGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	840
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCCGAGGCGCGGTATACATTGGGCGCCAGCTGTGTGCTGCTGCTGCC	900
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACCTACCTTTCTACGGAGTGGGATCTGCAACCTGCTGCTGCTGCTGCTGCTGCTG	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGCAGCCCTGTGCGCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGCTGTGGGCATGGAGCACTTCGGAGAGGTGAGGGCAGTTACAGTGCACAT	1080


```
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 1015 TCCATGCCCAATCCCGAGGCGGTATACATTCGGCGCAGCTGTGTGACTGCCTGTCC 1074
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1075 TACAACCTACCTTTCACGAGAGTGGATCCCTGACCCCTCGTGTGCCCTGCACAAACAA 1134
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1135 GAGGTGACAGCAGAGATGGACACAGCGGTGTGAGAAGTCAGCAAGCCCTGTGCCGA 1194
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1195 GTGTCTATGTCTGGCATGGAGCAGCTTGCAGAGGTGAGGGCAGTTTACCAGTGCCAAT 1254
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1255 ATCCAGGAGTTGTGGCTGCAAGAAGATCTTTGGGAGCCTGGCAATTTCTGCCGGAGAGC 1314
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1315 TTTGATGGGACCCAGCTCCAAACACTGCCCGCTCCAGCCAGACGACTCCAAGTGT 1374
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1375 GAGACTCTGGAGAGATCACAGGTATACCTATACATCTCAGCATGCGCGAGCAGCGCTGCT 1434
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1435 GACCTACCGTCTCCAGAACCTGCAAGTATCCGGGAGCAATTTCTGCACAAATGGCGCC 1494
QY 441 TyrSerLeuThrLeuGlnGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1495 TACTCGCTGACCTCGAAGGCTGGCATCAGCTGGCTGGGCTCGCTCTACTGAGGGAA 1554
QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1555 CTGGCAGTGACTGGCCCTCATCCACCATACACCCACCTCTGCTGCTGCACACGGTG 1614
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1615 CCCTGGGACCAAGCTCTTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCA 1674
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1675 GAGGACGAGTGTGTGGCGAGGCGCTGGCTGCCACCAAGCTGTGCGCCGAGGCGACTGC 1734
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1735 TGGGTCAGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC 1794
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1795 GTGGAGGAATGCCGAGTACTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1854
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1855 TTGCCGTGCCACCTGTAGTGTGACCCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1914
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1915 CTGACCAAGTGTGGGCTGTGCCCATATTAAGACCCCTCTCTGCGTGGCCGCTGC 1974
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1975 CCCAGCGTGTGAAACCTGACCTCTCTACATGCCCATCTCGAAGTTTCCAGATGAGGAG 2034
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 2035 GGCAGTATGCCAGCTTGCCCATCAACTGCACCCACTCTGTGTGGACCTGGATGACAAG 2094
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
```

```
Db 2095 GGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGTGACGTCC 2133
RESULT 13
ARI67390
LOCUS ARI67390 4473 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 26 from patent US 6287569.
ACCESSION ARI67390
VERSION ARI67390.1 GI:17903168
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4473)
AUTHORS Kipps T.J. and Wu Y.
TITLE Vaccines with enhanced intracellular processing
JOURNAL Patent: US 6287569-A 26 11-SEP-2001;
FEATURES
Location/Qualifiers
1..4473
source /organism="unknown"
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:
Pred. No.: 3.45e-226 Length: 4473
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ3 (1-653) x ARI67390 (1-4473)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProGlyAla 20
Db 175 ATGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCTCGCCCTTGTGCCCGGAGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCAAGTGTGCACCGGCACACATGAAGCTGCGCTCCCTGCCAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCCACTGGACATGCTCCGCCACTCTTACCAGGGCTGCCAGGTGGTGACGGAAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 355 GAATCCTACCTACCTGCCCAACCAATGCCAGCTGTCTTCTTCAGGAGATATCCAGGAGTG 414
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGCTACGTGCTCATCGCTCACAAACCAAGTAGGCGAGGTCCCACTGCAGAGGCTGCGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGTAGACAATGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGTGAACATACCAACCCCTGTACAGGGGCTCCCGAGGAGGCTCGGGAGGCTG 594
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTGAAGCCCTCACAGAGATCTTGAAGAGGGGTCTTGTATCCAGCGGAACCCCGAG 654
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 655 CTCTGCTACCAAGGACAGATTTTGTGAAGAGACATCTTCCAAAGAACCAACAGCAGTGCCT 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACACTGATAGACACCAACCGCTCTCGGGGCTGCCACCCCTGTTCTCCCATGTGTAG 774
```

QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 775 GGCCTCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTCAGAGCCTGACGCGCACTGCTGT 834
 QY 221 AlaGlyGlyCysAlaArgCysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 DB 835 GCGGTGGCTGTGCGCGCTGCAAGGGGCGCACATGCGCCACTGACTGCGCATGAGCAGTGT 894
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 895 GCTCGCGGCTGCACGGGCCCCAAGCAGCTCTGACTGCTGCGCTCCCTCCTCAACAC 954
 QY 261 SerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 955 AGTGGCATCTGTGAGCTGACCTGCCAGCGCTGCTCACCTACAACACAGACAGCTTGAG 1014
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 1015 TCCATGCCCAATCCCGAGGGCGGTATACATTGCGGCGCAGCTGTGACTGCTGCTGCC 1074
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 1075 TACAACCTACTTCTACGAGCGTGGGATCTGCACCTCGTCTGCCCGCTGCACACCAA 1134
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 1135 GAGGTACACAGAGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1194
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1195 GTGTGCTATGCTGGCGATGAGCAGCTTGCAGAGGTGAGGCGAGTTACCAAGTGCAT 1254
 QY 361 IleGlnIlePheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1255 ATCCAGGAGTTTGTGGCTGCAAGAGATCTTTGGAGCGCTGCATTTCTGCCGGAGAC 1314
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400
 DB 1315 TTTGTATGGGACCCAGCTCCCAACACTGCGCGCTCCAGCCAGAGCAGCTCCCAAGTGT 1374
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro 420
 DB 1375 GAGACTCTGGAGAGATCACAGTTTACCTATATACCTACATCTCAGATGCGCGGACAGCT 1434
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgTleLeuHisAsnGlyAla 440
 DB 1435 GACCTCAGGCTCTTCAGAACCTGCAGATATCCGGGAGCAATCTGCACATAGCGGCC 1494
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1495 TACTCGCTCACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTCAGTGGGAA 1554
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1555 CTGGGCGAGTGGAGCTGGCCCTCATCCACATAACACCCACTCTGCTGCTGCGACAGGTG 1614
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1615 CCTCGGACCACTCTTCGGAACCGCGACCAAGCTCTCTCATCACTGCCACCGGCCA 1674
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1675 GAGGAGAGTGTGTGGGCGAGGCGCTGGCGCTGCCACAGCTGTGGCGCCGAGGCGACTG 1734
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1735 TGGGGTCCAGGCGCCACCCAGTGTGTCAACTGCGAGCGAGTCTCTTCTGGGCGCAGGAGTGC 1794
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 DB 1795 GTGGAGGAATGCCAGTACTGACGGGCGCTCCCGAGGAGATGTGAATGCCAGGCACTGT 1854
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580

DB 1855 TTTGCCGTGCACCTGAGTGTGAGCCGCCAGAAATGGCTGAGTACCTGTTTGGACGGAG 1914
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAlaArgCys 600
 DB 1915 GCTGACCACTGTGTGCCCTGTGCCCCTATAGGACCTCCCTCTCTGCGTGGCCCCCTGC 1974
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 DB 1975 CCCAGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2034
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
 DB 2035 GCGCATGTCAGCCTTGGCCCATCAATGACCACTCTCTGTGTGACCTGGATGACAAG 2094
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 DB 2095 GGTGCCCCCGCAGCAGAGAGCCCTCTGACGTCC 2133
 RESULT 14
 HSERB2R
 LOCUS Human c-erb-B-2 mRNA. 4473 bp mRNA linear PRI 30-MAR-1995
 DEFINITION X03363
 ACCESSION
 VERSION
 KEYWORDS cell surface glycoprotein; cellular oncogene; erbB-2 cellular glycoprotein; growth factor receptor; kinase; neu cellular oncogene; transmembrane protein; tyrosine kinase.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 4473)
 AUTHORS Yamamoto,T., Ikawa,S., Akiyama,T., Semba,K., Nomura,N., Miyajima,N., Saito,T. and Toyoshima,K.
 TITLE Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor
 JOURNAL Nature 319 (6050), 230-234 (1986)
 MEDLINE 86118663
 PUBMED 3003577
 REFERENCE 2
 AUTHORS Papewalis,J., Nikitin,A.Yu. and Rajewsky,M.F.
 TITLE G to A polymorphism at amino acid codon 655 of the human erbB-2/HER2 gene
 JOURNAL Nucleic Acids Res. 19 (19), 5452 (1991)
 MEDLINE 92020265
 PUBMED 1681519
 COMMENT The c-erb-B-2 protein shows similarity to the epidermal growth factor receptor.
 FEATURES
 Location/Qualifiers
 1..4473
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /strain="cell line MKN-7"
 /db_xref="taxon:9606"
 175..3942
 /note="c-erb-B-2 precursor"
 /codon_start=1
 /protein_id="CAA27060.1"
 /db_xref="GI:31198"
 /translation="MELALCRWLLALLPGAASTQVCTGTDMLRUPASPTHLD MURHLYQGVQVQGNLELYLPTNASLFLQDIOEVQVYVLIAHNOVROPQRGRIV RGTQLFEDNYALVLDNGDPLANTPTVTGASPGGLRELRLSTELTKGVLIQRNPQ LCYQDTLWKDIFHKNNQLALTLIDNRSRACHPCSPMCKSGRCWGESSEDCOSLRT VACGACRCKGPLPTDCHEQCAAGCTGPKHSDCLACLFNHSIGTCLELCPALVTNT DTFESMPNPEGRYTFGASCATCPYNYLSTDVSGCKTLVCPLNQEVTAEDGTGQCEK SKPCARVCYGLGMEHLREVRVTSANIOEAGCKTIFGSLAFLPESFDGDPASNTAPL OPEQLQVETLEITGYLYISAMPDSLPDLSYFQNLQVIRGRILLINGAYSLTQIGLI SWGLRSLRELHSGGALIHNNHILCFVHTVPMDQLFRPNHQALLTANRPECVGEGL LACHOLCARHCHWGPQTQVNCQSFQVRGQCEVCRVQLQGLPRETYVNAHCLPCHPE CQPGNSVTFCFGEADQCACAKYKDPFCVAPCSGVKPKDLSYMPKWPPEEGACQ PCPINTHSCVDLDDKGCPAEQRASPLTIIISAVVGILLVVVGVVFGTILIKRROQKI

RRYTMRRLLOETELVEPLTPSGAMPNQAMRILKETELRKVKVLGSGAFGTYYKGIWI
 PDGENVKIPVAKVIRLNGTSKANKKEILDEAYVMAGSGSPVSVRLGLTCLSTVOLVT
 OLMYPGCLLDHVRNRRGKSQDLINMCQIAKMSYLEDVRLVHRDLAARNVLKSP
 NHVATDPLGLARLIDIDETHEYHADGGKVPKIMMALESILRRRFTHQSDVWSYGYTVRE
 LMTGAKYDGIIPAREIDPLEKGERLPQPPICHTIDVIMVKWMDISECRPRREL
 YSESRMARDPQFVVYIIONEDLGASPLDSTFYRSLLEDMDMDGLVDAEYLYVPOQGF
 FCPDAPAGAGMWHRHRSSTRGSGDLTLGLEPSEEEAPRSLAPSGAGSDVFDG
 DLGMAARKGLASPTHDPSPLQRYSEDPVLPSETDGYVAPLTCSPQPEYVNPQDVR
 PQQPSREGPLPAARPGATLERPKTLSPGKNGVVKOVAFEGGAIVENPEYLTPOGGAA
 PPHPPPAFSPAFNLVYWDQDPPERGAPPSTFKGTPTFAENPEYGLDVPV"

sig_peptide 175..237 /note="put. signal peptide (aa -21 to -1)"

mat_peptide 238..3939 /product="put. c-erb-B-2 protein (aa 1-1234)"

misc_feature 376..384 /note="pot. glycosylation site"

misc_feature 544..558 /note="pot. glycosylation site"

misc_feature 733..741 /note="pot. glycosylation site"

misc_feature 949..957 /note="pot. glycosylation site"

misc_feature 1762..1770 /note="pot. glycosylation site"

misc_feature 1885..1893 /note="pot. glycosylation site"

misc_feature 2059..2067 /note="pot. glycosylation site"

misc_feature 2353..3132 /note="aa 727-986, seq. homologous to EGF receptor kinase domain"

misc_feature 2446..2454 /note="pot. glycosylation site"

misc_feature 4455..4460 /note="put. polyA signal"

polyA_site 4473 /note="polyA site"

BASE COUNT 902 a 1383 c 1329 g 859 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,45e-226 Length: 4473
 Score: 3628.00 Matches: 653
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

SEQ3 (1-653) x HSERB2R (1-4473)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 DB 175 ATGGAGCTGGCGGCTTGTGGCGTGGGGGCTCTCTCGCCCTTGGCCCCCGGAGCC 234
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 235 GCGAGCACCACCAAGTGTGACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 294
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlnLeu 60
 DB 295 ACCCACCCTGGACATGCTCCCGCACCTCTACACAGGGCTGCCAGGTGGTGACAGGAAACCTG 354
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGlnVal 80
 DB 355 GAACTACCTACCTGCCACCAATGCCAGCTGCTCTCTCCCTGAGGATATCCAGAGGTG 414
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 415 CAGGCTACGTGCTCATCGCTCACAAACCAAGTAGGACAGGTCCTCCACTGCAGAGCTCGG 474
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 475 ATTTGGCAGGACCCAGCTCTTTGAGGACAACTATGCCCTTGGCGTGTCTAGACAAATGGA 534

121 AspProLeuAsnAsnThrThrProValThrGlyAlaLaserProGlyGlyLeuArgGluLeu 140
 DB 535 GACCCGCTGACAAATACACCCCTGTACAGGGGCTCCCCAGGAGGCTCGGGAGGTG 594
 QY 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuIleGlnArgAsnProGln 160
 DB 595 CAGCTTCAAGCCTCACAGAGATCTTGAAGAGAGGGGTCTTGATCCAGCGGAACCCCGAG 654
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 DB 655 CTCTGCTACCAAGGACACGATTTTGGGAAGGACATCTTCCACAGAACAACAGCTGGCT 714
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 715 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAAG 774
 QY 201 GlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 775 GGCTCCCGCTGCTGGGAGAGAGTCTGAGGATTTGTCAGAGCTGACGCGCACTGTCTGT 834
 QY 221 AlaGlyGlyCysAlaArgCysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 DB 835 GCCGTGGCTGTCCCGCTGCAAGGGCCACTGCCCACTGACTGCTGCTCCATGAGCAGTGT 894
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 895 GCTGCGGCTGCACGGGCCCCAAGCACTCTGACTGCTGCTGCTGCCCTCACTCAACCCAC 954
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 955 AGTGGCATCTGTGAGCTGACTGCCAGCCCTGTGCTACCTACAAACACAGACACGTTTGA 1014
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 1015 TCCATGCCCCAATCCGAGGGCCGTTATACATTCGGCGCCAGCTGTGTGACTGCTGCTGCC 1074
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 1075 TACAACCTACCTTCTACGGAGCTGGGATCCTGCACCCCTCGTGTGCCCTGCACAAACCA 1134
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 1135 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTCAGCAAGCCCTGTGCCGA 1194
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1195 GTGTGCTATGCTTGGGCATGAGCACCTTGCAGAGGTGAGGGAGTTTACCAGTCCCAAT 1254
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1255 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTTCCCGGAGAGC 1314
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB 1315 TTTGATGGGACCCAGCCTCCCAACTGCCCGCTCCAGCCAGAGACGCTCCAAGTGT 1374
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
 DB 1375 GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGCGCGGAGCAGCTGGCT 1434
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1435 GACTCAGCTGCTTCCAGAACCTGCAAGTATATCGGGGACCAATCTGCACAAATGCGCGC 1494
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1495 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGAA 1554
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1555 CTGGGCAGTGGAGCTGGCCCTCATCCACCATAACACCCCACTGTCTGCTGCTGCACAGGTG 1614

QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuLeuHisThrAlaAsnArgPro	500
Db	1615	CCCTGGGACCAAGCTTTTCGGAACCCGCCAACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1674
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1675	GAGGACGAGTGTGGGGGAGGGCTCGCCCTGCCACCAAGCTGTGGCCCGGAGGCACGTGC	1734
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1735	TGGGCTCCAGGCCCCACCACTGTCTCAACTGCAGCCAGTTCCTTCGGGGCCAGAGTGC	1794
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1795	GTGGAGGAATCCGAGTACTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACGTGT	1854
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1855	TTCGCGTCCCACTTGTAGTGTACGCCCCAGAAATGGCTCAGTACCTGTTTGGACCGGAG	1914
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1915	GCTGACCACTGTGGCCCTGTGCCACTATTAAGGACCTCCCTTCGTGCGCGCGCTGC	1974
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu	620
Db	1975	CCGAGCGTGTGAACCTGACCTCTCTTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2034
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2035	GGCGATGCCAGCCTTGCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAG	2094
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2095	GGCTGCCCGCGGACGACAGAGCCAGCCCTCTGAGGTCC	2133
RESULT 15			
LOCUS	AR202597	4530 bp	DNA linear PAT 20-JUN-2002
DEFINITION	Sequence 4 from patent US 6365151.		
ACCESSION	AR202597		
VERSION	AR202597.1 GI:21498768		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 4530)		
TITLE	Halpern,M.S. and England,J.M.		
JOURNAL	Cellular immunogens comprising cognate proto-oxogenes		
FEATURES	Patent: US 6365151-A 4 02-APR-2002;		
source	Location/Qualifiers		
	1..4530	/organism="unknown"	
BASE COUNT	922 a	1382 c	1346 g 880 t
ORIGIN			
Alignment Scores:			
Pred. No.:	3.5e-226	Length:	4530
Score:	3628.00	Matches:	653
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
SEQ3 (1-653) x AR202597 (1-4530)			
QY	1	MetGluLeuAlaLeuLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla	20
Db	151	ATGGAGCTGGCGGCTTGTGCGCTGGGGGCTCTCTCTCGCCCTCTTGCCTCCCGGAGCC	210
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	211	CGAGACCCCAAGTGTGCACCGGACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG	270

QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
Db	271	ACCCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGCTGCAGGGAACCTG	330
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	331	GAATCACTACCTGCTCCCAACCAATGCCAGCTGTCTTCTCTCGAGGATATCCAGAGGTG	390
QY	81	GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	391	CAGGCTACGTGCTCATCGTCACAAACCAAGTGAGGAGGTCCCACTGCAGAGGCTGCCG	450
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	451	ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGTGTAGACAATGGA	510
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	511	GACCGCTGAACAATACCACCCCTGTACAGGGGCTCCCAAGGAGGCTGCGGAGGTG	570
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Db	571	CAGCTTCAAGCCTCACAGAGATCTGAAAGAGGGGCTTTCATCCAGCGGAACCCCAAG	630
QY	161	LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	631	CTCTGCTACCAAGGACACCATTTTGTGAAGGACATCTTCCACAAGACAACCAAGCTGC	690
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	691	CTCACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAAG	750
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	751	GGCTCCCGCTGCTGGGGAGAGATTTCTGAGGATTGTACAGAGCTTACGGGCACTGTCTGT	810
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys	240
Db	811	GCCGGTGTCTGCCCGCTGCAAGGGCCACTGCCCACTGACTGTGCTGCATGAGCAGTGT	870
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	871	GCTCGCGCTGCACGGGCCCAAGCACCTCTGACTGCTGGCTGCCCTCCACTTCAACCCAC	930
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	931	AGTGGCATCTGTGAGCTGCATGCCCGCCCTGGTCACTCAACACACAGACAGCTTTAG	990
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	991	TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTCCC	1050
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1051	TACAATACCTTTCTACGAGCTGGGATCTCTGCACCCCTGCTGCTGCCCTTGCAACCAAA	1110
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1111	GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAGTGAGCAAGCCCTGTCCCCGA	1170
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1171	GTGTGCTATGCTTGGGCATGAGCACCTTTCGAGAGGTGAGGGCAGTTTACAGTGCCTAAT	1230
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1231	ATCCAGGAGTTTGTGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC	1290
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1291	TTTGATGGGACCCAGCCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGTGT	1350

```
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1351 GAGACTCTGGAGAGATACAGGTTACCTATACATCTCAGCATGCCGCGACAGCTGCCT 1410
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyVala 440
Db 1411 GACCTCAGCGTCTTCCAGAACCTGCAGTAATCCGGGAGCAATTCGCACAATGGCGCC 1470
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1471 TACTCGCTGACCTGCAAGGCGTGGCATCAGCTGGCTGGGCGTGCCTCACTCAGGGAA 1530
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1531 CTGGGCAGTGAGCTGGGCCCTCATCCACATAACACCCACCTCTGCTTCGTGCACACGGTG 1590
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1591 CCCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1650
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1651 GAGCAGAGTGTGTGGCGAGGGCGCTGGCCCTGCCACCACTGTGCGCCCGAGGGCACTGC 1710
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1711 TGGGGTCCAGGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC 1770
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1771 GTGGAGGAAATGCCAGTACTGCGAGGGCTCCCCAGGGAGTATGTAATGCCAGGCACTGT 1830
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1831 TTGCCGTGCCACCTGAGTGTACGCCCCAGAAATGGCTCACTGACCTGTTTGGACCGGAG 1890
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1891 GCTGACCAGTGTGTGGCTGTGCCACTATAAGGACCCCTCCCTTCGCGTGGCCGCGTGC 1950
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1951 CCCAGCGGTGTGAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 2011 GGGCATGCCAGCCTTCCCCCATCACTGCACCCACTCCTCTGTGTGGACCTGGATGACAAG 2070
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 2071 GGCTGCCCGCCGAGCAGAGAGCGGCCCTCTGACGTCC 2109
```

Search completed: October 16, 2003, 11:02:40
Job time : 9453.94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:47:19 : Search time 234.794 Seconds
(without alignments)
3105.423 Million cell updates/sec

Title: JUNC_SEQ8_SEQ5
Perfect score: 160
Sequence: 1 ERGCPAQRASPVTFQNEGLGPASPLDSTF 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/HOLLERAN480/runat_15102003_131913_20548/app_query.fasta_1.4685
-DB=EST_QFMT-fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -DOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGCN_1.12645 -runat_15102003_131913_20548 -NCPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST :
1: em_estba :
2: em_esthum :
3: em_estin :
4: em_estnu :
5: em_estov :
6: em_estpl :
7: em_estro :
8: em_htc :
9: gb_estl :
10: gb_est2 :
11: gb_htc :
12: gb_est3 :
13: gb_est4 :
14: gb_est5 :
15: em_estfun :
16: em_estom :
17: em_gss_hum :
18: em_gss_inv :
19: em_gss_pln :
20: em_gss_vrt :
21: em_gss_fun :
22: em_gss_mam :
23: em_gss_mus :
24: em_gss_pro :
25: em_gss_fod :
26: em_gss_phg :
27: em_gss_vrl :
28: gb_gss1 :

29: gb_gss2 :*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	83	51.9	278	10	BF923905	BF923905 QV4-NT025
2	83	51.9	294	10	BF757857	BF757857 CM4-CT057
3	83	51.9	337	10	BE091682	BE091682 IL2-BT073
c	83	51.9	348	9	AI909847	AI909847 QV-BT225-
5	83	51.9	378	14	CB266492	CB266492 1005398 H
6	83	51.9	423	12	BM703962	BM703962 UI-E-CK1-
7	83	51.9	464	2	HSM067221	Bx477782 Homo sapi
8	83	51.9	496	12	BM795191	BM795191 K-EST0076
9	83	51.9	505	12	BM854421	BM854421 K-EST0136
10	83	51.9	531	12	BM830170	BM830170 K-EST0103
11	83	51.9	536	14	CB129403	CB129403 K-EST0179
12	83	51.9	547	12	BM787824	BM787824 K-EST0066
13	83	51.9	567	2	HSM077766	Bx487579 Homo sapi
14	83	51.9	574	12	BM829991	BM829991 K-EST0103
15	83	51.9	630	2	HSM073298	Bx483130 Homo sapi
c	83	51.9	691	14	CB853376	CB853376 UI-CF-FNO
17	83	51.9	767	10	BF240297	BF240297 601905830
18	83	51.9	849	14	CA489799	CA489799 AGENCOURT
19	83	51.9	852	14	CD516283	CD516283 AGENCOURT
20	83	51.9	866	14	CA488343	CA488343 AGENCOURT
21	83	51.9	893	14	CA455141	CA455141 AGENCOURT
22	83	51.9	894	14	CA454131	CA454131 AGENCOURT
23	83	51.9	899	14	CD558386	CD558386 AGENCOURT
24	83	51.9	916	13	B0845369	B0845369 AGENCOURT
25	83	51.9	919	13	B0845391	B0845391 AGENCOURT
26	83	51.9	924	14	CA489100	CA489100 AGENCOURT
27	83	51.9	932	14	CA487981	CA487981 AGENCOURT
28	83	51.9	984	14	CA489084	CA489084 AGENCOURT
29	83	51.9	1004	14	CD515356	CD515356 AGENCOURT
30	83	51.9	4715	11	AF318349	AF318349 Homo sapi
31	81	50.6	220	9	AW763060	AW763060 uc60c11.y
32	81	50.6	377	9	AW415355	AW415355 49593 MAR
33	81	50.6	483	4	BX530001	BX530001 R2PD Mus
34	81	50.6	518	10	BF039952	BF039952 BP250015A
35	81	50.6	552	10	BF022141	BF022141 uy47c11.y
36	81	50.6	654	12	BI648646	BI648646 603277149
37	81	50.6	655	12	BI557778	BI557778 603236957
38	81	50.6	695	12	BI557170	BI557170 603239144
39	81	50.6	718	12	BI108229	BI108229 602902263
40	81	50.6	749	12	BI554781	BI554781 603236758
41	81	50.6	754	12	BI651822	BI651822 603299821
42	81	50.6	771	12	BI154609	BI154609 602904778
43	81	50.6	891	10	BF101641	BF101641 601753564
44	81	50.6	3110	11	AK031542	AK031542 Mus muscu
45	81	50.6	3372	11	BC046553	BC046553 Mus muscu

ALIGNMENTS

RESULT 1
BF923905
LOCUS QV4-NT0251-251100-599-all NT0251 Homo sapiens CDNA, mRNA sequence.
DEFINITION BF923905
ACCESSION BF923905
VERSION BF923905.1 GI:12319793
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 278)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovskij-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV4&t2=QV4-NT0251-251100-599-all&t3=2000-11-25&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 10

High quality sequence stop: 278.

Location/Qualifiers

source

1. .278

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="NT0251"

/note="Organ: nervous.tumor; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 52 a 78 c 79 g 69 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0896 Length: 278

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x BF923905 (1-278)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrphe 30

Db 135 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTC 188

RESULT 2

BF757857

LOCUS

CM4-CT0574-101100-428-f11

CT0574

Homo sapiens

linear

EST 12-JAN-2001

DEFINITION

BF757857

VERSION

BF757857.1

GI:12105757

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 294)

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV4&t2=QV4-NT0251-251100-599-all&t3=2000-11-25&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 10

High quality sequence stop: 278.

Location/Qualifiers

source

1. .278

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="NT0251"

/note="Organ: nervous.tumor; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 52 a 78 c 79 g 69 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0896 Length: 278

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x BF923905 (1-278)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrphe 30

Db 135 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTC 188

RESULT 2

BF757857

LOCUS

CM4-CT0574-101100-428-f11

CT0574

Homo sapiens

linear

EST 12-JAN-2001

DEFINITION

BF757857

VERSION

BF757857.1

GI:12105757

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 294)

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

JOURNAL

MEDLINE

PUBMED

COMMENT

10737800

CONTACT

Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&t2=CM4-CT0574-101100-428-f11&t3=2000-11-10&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 9

High quality sequence stop: 294.

Location/Qualifiers

source

1. .294

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="CT0574"

/note="Organ: colon; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 55 a 85 c 86 g 68 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0963 Length: 294

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x BF757857 (1-294)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrphe 30

Db 129 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTC 182

RESULT 3

BE091682/c

LOCUS

IL2-BT0731-240400-070-E05

BT0731

Homo sapiens

linear

EST 12-JUN-2000

DEFINITION

BE091682

ACCESSION

BE091682

VERSION

BE091682.1

GI:8482134

KEYWORDS

EST.

SOURCE

Homo sapiens

(human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 337)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE
20202663
PUBMED
COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-IL2-BT0731-240
400-070-E05st3-2000-04-24st4=1)
Seq primer: puc 18 forward
High quality sequence stop: 336.

FEATURES
source

Location/Qualifiers
1..337
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BT0731"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
, 716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 72 a 101 c 96 g 68 t
ORIGIN

Alignment Scores:

Pred. No.: 0.115 Length: 337
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x BE091682 (1-337)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 148 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 95

RESULT 4
AI909847

LOCUS CB266492 348 bp mRNA linear EST 30-MAR-2000
DEFINITION QV-BT225-050599-038 BT225 Homo sapiens cDNA, mRNA sequence.
ACCESSION AI909847
VERSION AI909847.1 GI:6500527
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 348)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.K., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

TITLE

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/seq/gethtml.pl?tl=QVst2-QV-BT225-038.html
st3=050599st4=1)
Seq primer: puc 18 forward.

FEATURES
source

Location/Qualifiers
1..348
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="female"
/dev_stage="Adult"
/clone_lib="BT225"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
, 716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 68 a 96 c 114 g 70 t
ORIGIN

Alignment Scores:

Pred. No.: 0.12 Length: 348
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 9 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AI909847 (1-348)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 15 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 68

RESULT 5
CB266492

LOCUS CB266492 378 bp mRNA linear EST 20-FEB-2003
DEFINITION 1005398 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
cDNA 5', mRNA sequence.
ACCESSION CB266492
VERSION CB266492.1 GI:28441078
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 378)
Yang,R.-Z., Shuldiner,A. and Gong,D.-W.
EST analysis of human adipose gene expression
Unpublished
Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition
University of Maryland
660 Redwood St, HH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgong@medicine.umaryland.edu
PCR Primers
FORWARD: CTCGGGAAGCGCCGCTTGTGTGGT
BACKWARD: AATACGACTACATAGCGGGAATTGG
Seq primer: GTTGGTACCCGGGAATTC.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES
source

Location/Qualifiers
1..378
/organism="Homo sapiens"
/mol_type="mRNA"

```

/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/notes="Vector: lambdaTriplex"
BASE COUNT 71 a 110 c 117 g 80 t
ORIGIN
Alignment Scores:
Pred. No.: 0.133 Length: 378
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservatives: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 14 Gaps: 0
JUNC_SEQ8_SEQ5 (1-30) x CB266492 (1-378)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
    ||| ||||||||||||||||||||||||||||||||||||||||||||
Db 124 GTGGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 177
RESULT 6
LOCUS BM703962
DEFINITION UI-E-CK1-afk-o-11-0-UI.r1 UI-E-CK1 Homo sapiens cDNA clone
VERSION BM703962
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 423)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8989548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: Bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1..423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="UI-E-CK1-afk-o-11-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK1"
/notes="Organ: eye; Vector: p7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-CK1 is a normalized cDNA library containing the
following tissue(s): Retina Foveal and Macular. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated

```

```

to an EcoR I adaptor, digested with Not I, and cloned
directionally into p7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (df)18 tail. The sequence tag for this
library is GTCC. This library was created for the program,
Gene Discovery in the Visual System, supported by National
Eye Institute (NEI).
BASE COUNT 78 a 129 c 131 g 85 t
ORIGIN
Alignment Scores:
Pred. No.: 0.154 Length: 423
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservatives: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0
JUNC_SEQ8_SEQ5 (1-30) x BM703962 (1-423)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
    ||| ||||||||||||||||||||||||||||||||||||||||||||
Db 147 GTGGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 200
RESULT 7
HSM067221
ID HSM067221 standard; RNA; EST; 464 BP.
XX AC BX477782;
XX SV BX477782.1
XX DT 09-MAY-2003 (Rel. 75, Created)
XX DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)
XX DE Homo sapiens mRNA; EST DKFZp686M12198_r1 (from clone DKFZp686M12198)
XX KW EST; expressed sequence tag.
XX OS Homo sapiens (human)
XX OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
XX OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX RN [1]
XX RP Bloecker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
XX RA Han M., Wiemann S.;
XX RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.
XX RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
XX CC This is the 5' sequence of the clone insert
XX CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
XX CC Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
XX CC sequenced by GBF (National Research Centre for Biotechnology
XX CC Ltd., Braunschweig/Germany) within the cDNA sequencing
XX CC consortium of the German Genome Project.
XX CC No sl sequence available.
XX CC This clone (DKFZp686M12198) is available at the RZPD in Berlin.
XX CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
XX CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX FH Key Location/Qualifiers
FH source 1..464
FT /db_xref="taxon:9606"
FT /mol_type="mRNA"
FT /organism="Homo sapiens"
FT /clone="DKFZp686M12198"
FT /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
FT DH10B; sites SfiIA + SfiIB"
FT /dev_stage="adult"

```

/clone="lid- szrwn0z0
/note="Organ: Stomach; Vector: pTZ18RPI; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

BASE COUNT	95 a	138 c	161 g
100	100	100	100
101	101	101	101
102	102	102	102
103	103	103	103
104	104	104	104
105	105	105	105
106	106	106	106
107	107	107	107
108	108	108	108
109	109	109	109
110	110	110	110
111	111	111	111
112	112	112	112
113	113	113	113
114	114	114	114
115	115	115	115
116	116	116	116
117	117	117	117
118	118	118	118
119	119	119	119
120	120	120	120
121	121	121	121
122	122	122	122
123	123	123	123
124	124	124	124
125	125	125	125
126	126	126	126
127	127	127	127
128	128	128	128
129	129	129	129
130	130	130	130
131	131	131	131
132	132	132	132
133	133	133	133
134	134	134	134
135	135	135	135
136	136	136	136
137	137	137	137
138	138	138	138
139	139	139	139
140	140	140	140
141	141	141	141
142	142	142	142
143	143	143	143
144	144	144	144
145	145	145	145
146	146	146	146
147	147	147	147
148	148	148	148
149	149	149	149
150	150	150	150
151	151	151	151
152	152	152	152
153	153	153	153
154	154	154	154
155	155	155	155
156	156	156	156
157	157	157	157
158	158	158	158
159	159	159	159
160	160	160	160
161	161	161	161
162	162	162	162
163	163	163	163
164	164	164	164
165	165	165	165
166	166	166	166
167	167	167	167
168	168	168	168
169	169	169	169
170	170	170	170
171	171	171	171
172	172	172	172
173	173	173	173
174	174	174	174
175	175	175	175
176	176	176	176
177	177	177	177
178	178	178	178
179	179	179	179
180	180	180	180
181	181	181	181
182	182	182	182
183	183	183	183
184	184	184	184
185	185	185	185
186	186	186	186
187	187	187	187
188	188	188	188
189	189	189	189
190	190	190	190
191	191	191	191
192	192	192	192
193	193	193	193

BASE COUNT

ORIGIN

Alignment Scores:
 Pred. No.: 0.193 Length: 505
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 12 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x BM854421 (1-505)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 Db 334 GTGGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 387
 RESULT 10
 BM830170 531 bp mRNA linear EST 06-MAR-2002
 DEFINITION K-EST0103592 S21SNU520s1 Homo sapiens cDNA clone S21SNU520s1-19-D03
 5', mRNA sequence.
 ACCESSION BM830170
 VERSION BM830170.1 GI:19186579
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 531)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.

21C Frontier Korean EST Project 2001

TITLE
JOURNAL
COMMENT

Unpublished
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@email.kribb.re.kr
 Plate: 19 row: D column: 03
 High quality sequence stop: 531.
 Location/Qualifiers

FEATURES

source

1. 531
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S21SNU520s1-19-D03"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="Floating aggregates"
 /cell_line="SNU-520"
 /lab_host="Top10r"
 /clone_lib="S21SNU520s1"

/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
 Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10r by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library. After analyzing and
 sequencing about 2,000 ~ 3,000 colonies in original cDNA
 library, the abundant cDNAs were selected and amplified by

PCR reaction using vector region primer including T7
 promoter as 5' primer and N(dt)14 as 3' primer. The PCR
 products were used as template for synthesis of
 biotinylated single stranded RNA by in vitro transcription
 reaction. The synthesized RNA probes were hybridized with
 antisense single stranded cDNAs prepared from original
 library and incubated with avidin-gel. After removing
 DNA-RNA hybrids by centrifuge, the subtracted cDNA
 libraries were constructed by transformation of the
 remaining DNA into competent cells E. coli Top10r with
 electroporation method."

BASE COUNT 102 a 149 c 166 g 114 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.206 Length: 531
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 12 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x BM830170 (1-531)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 334 GTGGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 387

RESULT 11

CB129403

LOCUS

DEFINITION

K-EST0179037 C1SNU17 Homo sapiens cDNA clone C1SNU17-5-C08 5', mRNA

sequence.

ACCESSION CB129403

VERSION CB129403.1 GI:28093000

KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 536)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.

21C Frontier Korean EST Project 2001

TITLE

JOURNAL

COMMENT

Unpublished

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@email.kribb.re.kr

Plate: 5 row: C column: 08

High quality sequence stop: 536.

Location/Qualifiers

1. 536

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="C1SNU17-5-C08"

/sex="F"

/tissue_type="Uterine"

/cell_type="Epithelial"

/cell_line="SNU-17"

/lab_host="Top10r"

/clone_lib="C1SNU17"

/note="Organ: Cervix; Vector: pCNS-D2; Site_1: EcoRI;
 Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including
 EcoRI site by treatment of T4 RNA ligase and the first

FT /clone="DKF2p686N09234"
FT /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
FT DH10B; sites SfiIA + SfiIB"
FT /dev_stage="adult"
FT /tissue_type="cDNA-collection"

XX
SQ Sequence 630 BP; 124 A; 174 C; 196 G; 133 T; 3 other;

Alignment Scores:
Pred. No.: 0.257 Length: 630
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x HSM073298 (1-630)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 439 GTGGTCATCCAGAAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACAGCACCTTC 492

Search completed: October 16, 2003, 17:04:07
Job time : 235.794 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 16, 2003, 11:08:07 ; Search time 31.0111 Seconds
(without alignments)
2540.503 Million cell updates/sec

Title: JUNC_SEQ8_SEQ5

Perfect score: 160

Sequence: 1 ERGCCPAERASPVTFQNEGLGASPILDSTF 30

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/HOLLERAN480/runat_15102003_131915_20662/app_query.fasta_1.4685
-DB=Published_Applications_NA -QPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USPR=HOLLERAN480 @CGN.1.1.397 @runat_15102003_131915_20662
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published_Applications_NA:
1: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2.6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2.6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2.6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2.6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2.6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2.6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2.6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2.6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2.6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2.6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2.6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2.6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2.6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2.6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2.6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2.6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description
1	83	51.9	1115	14	US-10-102-806-165 Sequence 165, App

2	83	51.9	1713	12	US-10-378-393-14 Sequence 14, Appl
3	83	51.9	1755	10	US-09-930-125-6 Sequence 6, Appli
4	83	51.9	1767	10	US-09-930-125-4 Sequence 4, Appli
5	83	51.9	1773	10	US-09-930-125-7 Sequence 7, Appli
6	83	51.9	1806	10	US-09-930-125-5 Sequence 5, Appli
7	83	51.9	2411	12	US-10-378-393-10 Sequence 10, Appl
8	83	51.9	3765	12	US-10-207-498-5 Sequence 5, Appli
9	83	51.9	3768	9	US-09-811-123-8 Sequence 8, Appli
10	83	51.9	3768	9	US-09-811-115-2 Sequence 9, Appli
11	83	51.9	3768	10	US-09-854-356-9 Sequence 1, Appli
12	83	51.9	3768	10	US-09-930-125-1 Sequence 1, Appli
13	83	51.9	3768	12	US-10-313-644-1 Sequence 1, Appli
14	83	51.9	4473	11	US-09-441-411-5 Sequence 5, Appli
15	83	51.9	4473	12	US-10-101-510-81 Sequence 81, Appl
16	83	51.9	4473	14	US-10-146-473-32 Sequence 32, Appl
17	83	51.9	4473	14	US-10-207-655-44 Sequence 44, Appl
18	83	51.9	4530	10	US-09-877-177-11 Sequence 11, Appl
19	83	51.9	4530	12	US-10-007-926A-119 Sequence 119, App
20	83	51.9	4530	12	US-10-101-510-124 Sequence 124, App
21	83	51.9	4530	12	US-10-338-730-1 Sequence 1, Appli
22	83	51.9	4530	14	US-10-177-293-125 Sequence 125, App
23	83	51.9	4543	10	US-09-769-508-1 Sequence 1, Appli
24	83	51.9	4606	12	US-09-971-392-70 Sequence 70, Appl
25	83	51.9	4642	14	US-10-198-846-10896 Sequence 10896, A
26	83	51.9	9274	9	US-09-811-123-7 Sequence 7, Appli
27	83	51.9	9274	9	US-09-811-115-1 Sequence 1, Appli
28	81	50.6	3771	10	US-09-854-356-11 Sequence 11, Appl
29	81	50.6	3955	10	US-09-870-759-117 Sequence 117, App
30	81	50.6	3955	10	US-09-854-356-10 Sequence 10, Appl
31	81	50.6	3955	12	US-09-751-708A-117 Sequence 117, App
32	66	41.2	201	13	US-10-109-213-3 Sequence 3, Appli
33	58.5	36.6	143601	12	US-09-855-824-3 Sequence 25, Appl
34	58	36.2	1938	12	US-10-422-264-25 Sequence 27, Appl
35	58	36.2	2316	12	US-10-422-264-27 Sequence 23, Appl
36	58	36.2	2604	12	US-10-422-264-23 Sequence 23, Appl
37	58	36.2	3320	10	US-09-862-658-1 Sequence 1, Appli
38	58	36.2	3320	14	US-10-175-696-22 Sequence 22, Appl
39	58	36.2	3384	12	US-10-422-264-29 Sequence 29, Appl
40	57.5	35.9	363	11	US-09-918-995-28706 Sequence 28706, A
C 41	57	35.6	23432	9	US-09-764-869-1332 Sequence 1332, Ap
C 42	57	35.6	23432	14	US-10-091-504-1332 Sequence 1332, Ap
C 43	56.5	35.3	1557	10	US-09-774-434-6 Sequence 6, Appli
C 44	56.5	35.3	22452	10	US-09-764-868-1487 Sequence 1487, Ap
C 45	56.5	35.3	22452	10	US-09-764-868-1489 Sequence 1489, Ap

ALIGNMENTS

RESULT 1

US-10-102-806-165
; Sequence 165, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103PIC1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 165
; LENGTH: 1115
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (390)

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-165

Alignment Scores: Length: 1115
Pred. No.: 0.00141 83.00
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservatives: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 14 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-10-102-806-165 (1-1115)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| |||
Db 446 GTGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACACGACCTTC 499

RESULT 2

US-10-378-393-14
; Sequence 14, Application US/10378393
; Publication No. US20030182669A1
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; CURRENT FILING DATE: 2003-03-03
; PRIOR FILING DATE: 2003-03-03
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-14

Alignment Scores: Length: 1713
Pred. No.: 0.00225 83.00
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservatives: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-10-378-393-14 (1-1713)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| |||
Db 904 GTGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACACGACCTTC 957

RESULT 3

US-09-930-125-6
; Sequence 6, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: OF HER-2/NEU-ASSOCIATED MALIGNANCIES

; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-6

Alignment Scores: Length: 1755
Pred. No.: 0.00231 83.00
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservatives: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-930-125-6 (1-1755)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| |||
Db 937 GTGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACACGACCTTC 990

RESULT 4

US-09-930-125-4
; Sequence 4, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-4

Alignment Scores: Length: 1767
Pred. No.: 0.00233 83.00
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservatives: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-930-125-4 (1-1767)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| |||
Db 937 GTGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACACGACCTTC 990

RESULT 5

US-09-930-125-7
; Sequence 7, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.

```
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-7

Alignment Scores:
Pred. No.: 0.00233 Length: 1773
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-930-125-7 (1-1773)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 961 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1014

RESULT 6
US-09-930-125-5
; Sequence 5, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-5

Alignment Scores:
Pred. No.: 0.00238 Length: 1806
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-930-125-5 (1-1806)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 994 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1047

RESULT 7
US-10-378-393-10
; Sequence 10, Application US/10378393
; Publication No. US20030182668A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVAT
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-10

Alignment Scores:
Pred. No.: 0.00327 Length: 2411
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-10-378-393-10 (1-2411)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 1602 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1655

RESULT 8
US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HERGULIN AND HER3
; FILE REFERENCE: 30448.103-US-01
; CURRENT APPLICATION NUMBER: US/10/207,498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5

Alignment Scores:
Pred. No.: 0.00532 Length: 3765
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-10-207-498-5 (1-3765)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
```

```

Db      2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACACACCTTC 3012
|||||
RESULT 9
US-09-811-123-8
; Sequence 8, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811.123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-123-8

Alignment Scores:
Pred. No.:      0.00532      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    51.88%      Indels:      0
DB:              9      Gaps:      0

JUNC_SEQ8_SEQ5 (1-30) x US-09-811-123-8 (1-3768)

Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db      2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACACACCTTC 3012
|||||
RESULT 10
US-09-811-115-2
; Sequence 2, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811.115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred. No.:      0.00532      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    51.88%      Indels:      0
DB:              9      Gaps:      0

JUNC_SEQ8_SEQ5 (1-30) x US-09-811-115-2 (1-3768)

```

```

Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db      2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACACACCTTC 3012
|||||
RESULT 11
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc.feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.:      0.00532      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    51.88%      Indels:      0
DB:              10      Gaps:      0

JUNC_SEQ8_SEQ5 (1-30) x US-09-854-356-9 (1-3768)

Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db      2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACACACCTTC 3012
|||||
RESULT 12
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125

; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1

; LENGTH: 3768
; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: CDS
; LOCATION: (1)...(3765)

US-09-930-125-1

Alignment Scores:
Pred. No.: 0.00532 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-930-125-1 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 13

US-10-313-644-1

; Sequence 1, Application US/10313644
; Publication No. US20030157119A1

; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.

; TITLE OF INVENTION: HAND-ZIMMERMAN, SUSAN

; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; FILE REFERENCE: 210121.493C3

; CURRENT APPLICATION NUMBER: US/10/313,644
; CURRENT FILING DATE: 2002-12-04

; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 3768
; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: CDS
; LOCATION: (1)...(3765)

US-10-313-644-1

Alignment Scores:
Pred. No.: 0.00532 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-10-313-644-1 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 14

US-09-441-411-5

; Sequence 5, Application US/09441411
; Publication No. US20030008342A1

; GENERAL INFORMATION:

; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.

; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik

; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409

; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16

; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5
; LENGTH: 4473

; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-441-411-5

Alignment Scores:
Pred. No.: 0.00642 Length: 4473
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 11 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-441-411-5 (1-4473)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 3133 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3186

RESULT 15

US-10-101-510-81

; Sequence 81, Application US/10101510
; Publication No. US20030148295A1

; GENERAL INFORMATION:

; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN

; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE

; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510

; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947

; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 81

; LENGTH: 4473
; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-101-510-81

Alignment Scores:
Pred. No.: 0.00642 Length: 4473
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-10-101-510-81 (1-4473)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 3133 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3186

Search completed: October 17, 2003, 03:53:14
Job time : 34.0111 secs

THIS PAGE BLANK (USPTO)


```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 0.00745 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-08-625-101-1 (1-3768)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 3012
RESULT 2
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3768
; OTHER INFORMATION: /note= "product = "cerB-b2""
```

```
US-08-356-786-1

Alignment Scores:
Pred. No.: 0.00745 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-08-356-786-1 (1-3768)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 3012
RESULT 3
US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; TITLE OF INVENTION: HUMAN HER-2 EXPRESSION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748ris LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048,804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-5100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: NO
US-09-048-804-1

Alignment Scores:
Pred. No.: 0.0092 Length: 4473
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-048-804-1 (1-4473)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3133 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 3186
RESULT 4
US-09-056-105-26
```

```
; Sequence 26, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
; APPLICANT: KIPPS, THOMAS J.
; APPLICANT: WU, YUNQI
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
; FILE REFERENCE: 233/221
; CURRENT FILING DATE: 1997-04-10
; EARLIER APPLICATION NUMBER: 50/043,467
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-056-105-26

Alignment Scores:
Pred. No.: 0.0092 Length: 4473
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-056-105-26 (1-4473)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3133 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3186

RESULT 5
US-08-229-515A-9
; Sequence 9, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIYUDDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-9880
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-08-229-515A

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 6
US-08-645-865-9
; Sequence 9, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIYUDDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-9880
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-645-865-9

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-08-645-865-9 (1-4530)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 7
US-09-167-322-4
```

```
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-167-322-4 (1-4530)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 8
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
US-09-527-487-1

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-527-487-1 (1-4530)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 9
US-09-877-177A-11
; Sequence 11, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-877-177A-11 (1-4530)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 10
US-08-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-229-515A-14

Alignment Scores:
Pred. No.: 0.0158 Length: 3955
Score: 81.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.62% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-08-229-515A-14 (1-3955)
Qy 1 GluArgGlyCysProAlaGlnArgAlaSerProValThrPhe 15
Db 1946 GAACGAGGCTGCCAGCAGACGAGAGCCGCCGCGTGACATTC 1990

RESULT 11
US-08-865-14
; Sequence 14, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIYUDDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; US-08-645-865-14

Alignment Scores:
Pred. No.: 0.0158 Length: 3955
Score: 81.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.62% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-08-645-865-14 (1-3955)
Qy 1 GluArgGlyCysProAlaGlnArgAlaSerProValThrPhe 15
Db 1946 GAACGAGGCTGCCAGCAGACGAGAGCCGCCGCGTGACATTC 1990

RESULT 12
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Denlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; US-09-146-283-3

Alignment Scores:
Pred. No.: 0.766 Length: 2385
Score: 68.00 Matches: 17
Percent Similarity: 66.67% Conservative: 3
Best Local Similarity: 56.67% Mismatches: 6
Query Match: 42.50% Indels: 4
DB: 2 Gaps: 2

JUNC_SEQ8_SEQ5 (1-30) x US-09-146-283-3 (1-2385)

```



```

; Patent NO. 6025340
; GENERAL INFORMATION:
; APPLICANT: Springer, Caroline J
; APPLICANT: Marais, Richard
; TITLE OF INVENTION: Surface expression of enzyme in gene directed prodrug therapy
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6025340th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,251
; FILING DATE: 31-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/01782
; FILING DATE: 27-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9415167.7
; FILING DATE: 27-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 620-20
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 153 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-776-251-3

Alignment Scores:
Pred. No.: 0.0516 Length: 153
Score: 66.00 Matches: 11
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 78.57% Mismatches: 0
Query Match: 41.25% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-08-776-251-3 (1-153)

Qy 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThr 14
Db 10 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGAGC 51

```

Search completed: October 16, 2003, 17:16:45
Job time : 16.6049 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 28.4941 Seconds
(without alignments)
2842.104 Million cell updates/sec

Title: JUNC_SEQ8_SEQ5

Perfect score: 160
Sequence: 1 ERGCPAEQASPVTFQNEGLGPASPLDSTF 30

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-framer-p2n.model -DEV-xlp
-Q/cn2_1/USPTO_spoal_p/HOLLERAN480/runat_15102003_131912_20526/app_query.fasta_1.4695
-DB=N_Geneseq_19Jun03 -OFAST-fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGCN 1.1 1758 -runat_15102003_131912_20526 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	147	91.9	2763	24	ABA92252	Mouse Her-2/neu ex
2	147	91.9	2781	24	ABA92253	Her-2/neu extracel
3	83	51.9	1115	21	AAF21778	Human breast and o
4	83	51.9	1755	24	AAD32746	Human cDNA for the
5	83	51.9	1767	24	AAD32744	Human cDNA for the
6	83	51.9	1773	24	AAD32747	Human cDNA for the
7	83	51.9	1806	24	AAD32745	Human cDNA for the
8	83	51.9	3600	21	AAA89736	Human HER-2/neu co
9	83	51.9	3678	24	ABK86207	cDNA encoding huma
10	83	51.9	3768	17	AAT40739	HER-2/neu oncogene
11	83	51.9	3768	20	AAO1912	Human HER-2/neu on
12	83	51.9	3768	21	AAA09455	Human heregulin 2
13	83	51.9	3768	22	AAH23392	Human HER-2/neu pr
14	83	51.9	3768	24	ABZ35744	Human ERBB2 polynu
15	83	51.9	3768	24	ABX0987	Human ERBB2 DNA fr
16	83	51.9	3768	24	AAD43935	Human ERBB2 cDNA.
17	83	51.9	3768	24	ABV78168	Human ERBB2 DNA SE
18	83	51.9	3768	24	ABV78168	Human ERBB2 DNA SE
19	83	51.9	3768	24	AAD32743	Human Her-2/neu pr
20	83	51.9	3768	24	ABA92250	Human Her-2/neu CD
21	83	51.9	3768	24	ABK10730	Human Her-2/neu DN
22	83	51.9	3768	24	ABL91709	Human polynucleoti
23	83	51.9	3768	24	ABK14058	Human HER2 (Erbb2)
24	83	51.9	4299	14	AAQ46083	Sequence encoding the
25	83	51.9	4472	21	AA14812	cDNA encoding the
26	83	51.9	4473	19	ASQ76220	Human tumour antiq
27	83	51.9	4473	20	AAZ31071	HER-2 nucleic acid
28	83	51.9	4473	24	ABZ34969	Human gene express
29	83	51.9	4473	24	AAD38904	Human Her-2 DNA.
30	83	51.9	4530	16	AAT01585	Her-2/neu (ERBB2/c
31	83	51.9	4530	18	AAT71253	Human ER2 gene.
32	83	51.9	4530	21	AAZ60815	Nucleotide sequenc
33	83	51.9	4530	22	AAD19731	Human tyrosine kin
34	83	51.9	4530	24	ABV94128	Human gene express
35	83	51.9	4530	24	ABV94128	Breast carcinoma r
36	83	51.9	4530	24	ABN85585	Human HER2-neu SQ
37	83	51.9	4530	24	ABK83918	Human cDNA differe
38	83	51.9	4530	25	ACC50139	Breast cancer asso
39	83	51.9	4530	25	ABQ83856	Human Her2/Neu enc
40	83	51.9	9274	22	AAF24297	HER2 transgene pla
41	83	51.9	9274	24	AAD43934	HER-2 transgene pl
42	83	51.9	9274	24	ABK14057	Human HER2 (Erbb2)
43	81	50.6	3771	21	AAA89737	Mouse Her-2/neu CD
44	81	50.6	3771	22	AAH42210	Nucleotide sequenc
45	81	50.6	3771	24	ABA92251	Mouse Her-2/neu CD

ALIGNMENTS

RESULT 1
ABA92252
ID ABA92252 standard; cDNA; 2763 BP.
XX
AC ABA92252;
XX
DT 17-JUN-2002 (first entry)
XX
DE Mouse Her-2/neu extracellular-phosphorylation domain fusion cDNA.
XX
KW Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
XX
XX receptor; mouse; gene therapy; gene; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 1..2763

```

FT      /*tag= a
FT      /product= "Her-2/neu ECD-PD fusion"
XX
XX
XX
PN      WO200212341-A2.
XX
XX
PD      14-FEB-2002.
XX
XX
PF      03-AUG-2001; 2001WO-US24283.
XX
XX
PR      03-AUG-2000; 2000US-0632507.
XX
XX
PA      (CORI-) CORIXA CORP.
PA      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX
PI      Cheever MA, Gheysen D;
XX
XX
DR      WPI; 2002-241743/29.
DR      P-PSDB; AAM51152.
XX
XX
PT      Her-2/neu fusion protein for treating or preventing cancer by eliciting
PT      or enhancing an immune response to the protein, has Her-2/neu
PT      extracellular domain fused to Her-2/neu intracellular or
PT      phosphorylation domain
XX
XX
PS      Disclosure; Fig 23; 141pp; English.
XX
XX
CC      The present sequence is that of cDNA encoding a fusion between
CC      the extracellular domain (ECD) and the phosphorylation domain (PD)
CC      murine Her-2/neu oncoprotein (see AAM51151). The Her-2/neu gene is
CC      amplified and the oncoprotein is overexpressed in a variety of
CC      human cancers, including breast, ovarian, colon, lung and prostate
CC      cancer. Her-2/neu overexpression correlates with a poor prognosis
CC      in breast and ovarian cancers. The invention provides Her-2/neu
CC      fusion proteins, nucleic acids encoding them, viral vectors, and
CC      vaccines comprising the fusion proteins or nucleic acid molecules.
CC      In preferred fusion proteins, the ECD of Her-2/neu is fused to a
CC      Her-2/neu intracellular domain or PD (or its DeltapD fragment). An
CC      immune response to Her-2/neu protein is elicited or enhanced by
CC      administering the fusion protein in the form of a vaccine, or by
CC      transfecting cells of an animal ex vivo with a nucleic acid
CC      encoding the fusion protein, and delivering the transfected cells
CC      to the animal. The fusion proteins, nucleic acids, and isolated
CC      specific T-cells are useful for inhibiting the development of a
CC      cancer, especially breast, ovarian, colon, lung or prostate cancer
CC      in a patient. T cells that specifically react with a Her-2/neu
CC      fusion protein can be used to remove tumour cells from a sample in
CC      order to inhibit the development of cancer in a patient.
XX
XX
SQ      Sequence 2763 BP; 571 A; 855 C; 772 G; 565 T; 0 other;

Alignment Scores:
Pred. No.:      1.96e-11      Length:      2763
Score:          147.00      Matches:      27
Percent Similarity: 96.67%      Conservative: 2
Best Local Similarity: 90.00%      Mismatches: 1
Query Match:     91.88%      Indels:      0
DB:             24      Gaps:        0

JUNC_SEQ8_SEQ5 (1-30) x ABA92252 (1-2763)

Qy      1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db      1918 GAACGAGGCTGCCAGCAGAGCAGAGAGCCAGCCAGTGCAGTCTCAGACGAGGACTTA 1977
Qy      21 GlyProAlaSerProLeuAspSerThrPhe 30
Db      1978 GGCCCCCTCCAGCCATCGAGCAGACCTTC 2007

RESULT 2
ID      ABA92253
XX      ABA92253 standard; cDNA; 2781 BP.
AC      ABA92253;

```

```

XX      17-JUN-2002 (first entry)
XX
XX      Her-2/neu extracellular-phosphorylation domain-TcP0 fusion cDNA.
DE
XX
XX
KW      Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
KW      receptor; TcP0; mouse; gene therapy; gene; ss.
XX
XX      Chimeric - Mus musculus.
OS      Chimeric - Unidentified.
XX
XX      Key      Location/Qualifiers
FT      CDS      1..2781
FT      FT      /*tag= a
FT      FT      /product= "Her-2/neu ECD-PD-TcP0 fusion"
XX
XX      WO200212341-A2.
XX
XX
PD      14-FEB-2002.
XX
XX
PF      03-AUG-2001; 2001WO-US24283.
XX
XX
PR      03-AUG-2000; 2000US-0632507.
XX
XX
PA      (CORI-) CORIXA CORP.
PA      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX
PI      Cheever MA, Gheysen D;
XX
XX
DR      WPI; 2002-241743/29.
DR      P-PSDB; AAM51153.
XX
XX
PT      Her-2/neu fusion protein for treating or preventing cancer by eliciting
PT      or enhancing an immune response to the protein, has Her-2/neu
PT      extracellular domain fused to Her-2/neu intracellular or
PT      phosphorylation domain
XX
XX
PS      Disclosure; Fig 25; 141pp; English.
XX
XX
CC      The present sequence is that of cDNA encoding a fusion between
CC      the extracellular domain (ECD) and the phosphorylation domain (PD)
CC      murine Her-2/neu oncoprotein (see AAM51151) plus a C-terminal
CC      TcP0 motif that improves immunogenicity. The Her-2/neu gene is
CC      amplified and the oncoprotein is overexpressed in a variety of
CC      human cancers, including breast, ovarian, colon, lung and prostate
CC      cancer. Her-2/neu overexpression correlates with a poor prognosis
CC      in breast and ovarian cancers. The invention provides Her-2/neu
CC      fusion proteins, nucleic acids encoding them, viral vectors, and
CC      vaccines comprising the fusion proteins or nucleic acid molecules.
CC      In preferred fusion proteins, the ECD of Her-2/neu is fused to a
CC      Her-2/neu intracellular domain or PD (or its DeltapD fragment). An
CC      immune response to Her-2/neu protein is elicited or enhanced by
CC      administering the fusion protein in the form of a vaccine, or by
CC      transfecting cells of an animal ex vivo with a nucleic acid
CC      encoding the fusion protein, and delivering the transfected cells
CC      to the animal. The fusion proteins, nucleic acids, and isolated
CC      specific T-cells are useful for inhibiting the development of a
CC      cancer, especially breast, ovarian, colon, lung or prostate cancer
CC      in a patient. T cells that specifically react with a Her-2/neu
CC      fusion protein can be used to remove tumour cells from a sample in
CC      order to inhibit the development of cancer in a patient.
XX
XX
SQ      Sequence 2781 BP; 574 A; 859 C; 779 G; 569 T; 0 other;

Alignment Scores:
Pred. No.:      1.97e-11      Length:      2781
Score:          147.00      Matches:      27
Percent Similarity: 96.67%      Conservative: 2
Best Local Similarity: 90.00%      Mismatches: 1
Query Match:     91.88%      Indels:      0
DB:             24      Gaps:        0

JUNC_SEQ8_SEQ5 (1-30) x ABA92253 (1-2781)

```

QY 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
 Db 1918 GAACGAGGCTGCCAGCAGACAGAGAGCCAGCCAGTGACGCTCTCAGACGAGACTTA 1977

QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
 Db 1978 GGCCCTCCAGCCCATGAGCAGCACCTTC 2007

RESULT 3
 AAF21778
 ID AAF21778 standard; DNA; 1115 BP.
 XX
 AC AAF21778;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Human breast and ovarian cancer associated antigen gene SEQ ID 165.
 XX
 KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW nontropic; neurprotection; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200055173-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05881.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-611515/58.
 DR P-PSDB; AAB58875.
 XX
 PT New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases -
 XX
 PS Claim 1; Page 604; 1299pp; English.
 XX
 CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC nontropic; neurprotection; antiviral; antiallergic; hepatotropic;
 CC antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;
 CC antibacterial; antifungal; antiparasitic and cardiant activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemias; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 XX
 SQ Sequence 1115 BP; 210 A; 336 C; 338 G; 222 T; 9 other;

Alignment Scores:
 Pred. No.: 0.0139 Length: 1115
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 21 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AAF21778 (1-1115)
 QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 Db 446 GTGGTCATCCAGAAATGAGGACTTGGCCAGCCAGCTCCCTTGGACAGCACCTTC 499

RESULT 4
 AAD32746
 ID AAD32746 standard; cDNA; 1755 BP.
 XX
 AC AAD32746;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Human cDNA for the clone HICD_native_coding_region.
 XX
 KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1755
 FT /*tag= a
 FT /product= "Human protein encoded by cDNA for the clone
 FT HICD_native_coding_region"
 FT /transl_except= (pos:1741..1752, as:Leu-Glu)
 FT /note= "CDS does not include stop codon"
 FT /partial
 XX
 PN WO200214503-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 14-AUG-2001; 2001WO-US41733.
 XX
 PR 14-AUG-2000; 2000US-225152P.
 PR 28-SEP-2000; 2000US-236428P.
 PR 21-FEB-2001; 2001US-270520P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
 PI McNeill PD, Vedvick TS;
 XX
 DR WPI; 2002-280758/32.
 DR P-PSDB; AAE20483.
 XX
 PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,
 PT prevention and diagnosis of cancer, preferably breast cancer -
 XX
 PS Example 5; Page 118-119; 129pp; English.
 XX
 CC The invention relates to an isolated Her-2/Neu polypeptide composition
 CC effective for eliciting an immune response. The invention is useful for
 CC eliciting an immune response in a patient, where the patient is human
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
 CC The composition is useful for the therapy and diagnosis of cancer,
 CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
 CC and other compositions for the diagnosis, prevention and treatment of
 CC human malignancies, for stimulating and/or expanding T cells specific for
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
 CC patient. The invention is useful for stimulating a T cell response in a
 CC human patient, as probe or primer for nucleic acid hybridisation, to

XX Example 5; Page 119; 129pp; English.

PS The invention relates to an isolated Her-2/Neu polypeptide composition

CC effective for eliciting an immune response. The invention is useful for

CC eliciting an immune response in a patient, where the patient is human

CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.

CC The composition is useful for the therapy and diagnosis of cancer,

CC and other compositions, in pharmaceutical compositions, e.g., vaccine

CC human malignancies, for stimulating and/or expanding T cells specific for

CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a

CC patient. The invention is useful for stimulating a T cell response in a

CC human patient, as probe or primer for nucleic acid hybridisation, to

CC selectively form duplex molecules with complementary stretches of the

CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full

CC length gene from a suitable library, and to direct expression of a

CC polypeptide in appropriate host cells. The composition is useful in

CC prophylactic or therapeutic applications and for the treatment of cancer,

CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-

CC associated malignancies. The invention is useful in gene therapy. The

CC present sequence is human cDNA for the clone HICD_in_pPDM_coding_

XX sequence.

XX Sequence 1773 BP; 383 A; 528 C; 530 G; 332 T; 0 other;

SQ

Alignment Scores:

Pred. No.: 0.0242 Length: 1773

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 24 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AAD32747 (1-1773)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 961 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 1014

RESULT 7

AAD32745

ID AAD32745 standard; cDNA; 1806 BP.

AC AAD32745;

DT 01-JUL-2002 (first entry)

XX Human cDNA for the clone HICD_plus_8_HIS.

DE Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;

KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1803

FT /*tag= a

FT /product= "Human protein encoded by cDNA for the clone

FT HICD_plus_8_HIS"

FT /transl_except= (pos:1543..1545, aa:Pro)

XX WO200214503-A2.

XX 21-FEB-2002.

XX 14-AUG-2001; 2001WO-US41733.

XX 14-AUG-2000; 2000US-225152P.

PR 28-SEP-2000; 2000US-236428P.

PR 21-FEB-2001; 2001US-270520P.

XX (CORI-) CORIXA CORP.

XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;

PI McNeill PD, Vedvick TS;

XX WPI; 2002-280758/32.

DR P-PSDB; AAE20482.

XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,

PT prevention and diagnosis of cancer, preferably breast cancer

XX Example 5; Page 118; 129pp; English.

CC The invention relates to an isolated Her-2/Neu polypeptide composition

CC effective for eliciting an immune response. The invention is useful for

CC eliciting an immune response in a patient, where the patient is human

CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.

CC The composition is useful for the therapy and diagnosis of cancer,

CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine

CC and other compositions, for stimulating and/or expanding T cells specific for

CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a

CC patient. The invention is useful for stimulating a T cell response in a

CC human patient, as probe or primer for nucleic acid hybridisation, to

CC selectively form duplex molecules with complementary stretches of the

CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full

CC length gene from a suitable library, and to direct expression of a

CC polypeptide in appropriate host cells. The composition is useful in

CC prophylactic or therapeutic applications and for the treatment of cancer,

CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-

CC associated malignancies. The invention is useful in gene therapy. The

CC present sequence is human cDNA for the clone HICD_plus_8_HIS.

XX sequence.

XX Sequence 1806 BP; 391 A; 530 C; 544 G; 341 T; 0 other;

SQ

Alignment Scores:

Pred. No.: 0.0248 Length: 1806

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 24 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AAD32745 (1-1806)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 994 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 1047

RESULT 8

AAA89736

ID AAA89736 standard; DNA; 3600 BP.

XX AAA89736;

XX 12-JAN-2001 (first entry)

DT Human HER-2/neu coding sequence.

DE Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;

KW breast cancer; prostate cancer; ovarian cancer; lung cancer;

KW colon cancer; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..3600

FT /*tag= a

FT /product= "HER-2/neu protein"

XX WO2000044899-A1.

XX 03-AUG-2000.

XX

PF 28-JAN-2000; 2000WO-US02164.
XX
PR 29-JAN-1999; 99US-0117976.
XX
XX (CORI-) CORIXA CORP.
PA (SMIK) SMITHKLINE BEECHAM.
XX
XX Cheever MA, Gheysen D;
PI
XX WPI; 2000-505976/45.
DR P-PSDB; AAB21198, AAB21208.
DR
XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins
XX useful for vaccinating against breast, ovarian, colon, lung and
XX prostate cancers
XX
XX Disclosure; Fig 15; 128pp; English.
XX
XX The present sequence encodes the human HER-2/neu protein. HER-2/neu is
CC a member of the tyrosine kinase family of receptor-like glycoproteins
CC and shows homology to the epidermal growth factor receptor (EGFR). It
CC probably plays a part in cell growth and/or differentiation. The
CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising
CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
XX
XX Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 0.0566 Length: 3600
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 21 Gaps: 0
JUNC_SEQ8_SEQ5 (1-30) x AAB89736 (1-3600)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACACACCTTC 3012
RESULT 9
ABK86207
ID ABR86207 standard; cDNA; 3678 BP.
XX
AC ABR86207;
XX
XX 24-SEP-2002 (first entry)
XX
XX cDNA encoding human breast cancer antigen, Her2 variant.
DE
XX Human; Her2; cytostatic; antiviral; immunostimulant;
XX cell-mediated immune response; tumour; breast cancer;
KW virus infection; prostate cancer; colorectal cancer; pancreatic cancer;
KW lymphoma; leukaemia; hepatitisvirus; lentivirus; herpesvirus;
KW human immunodeficiency virus; HIV; flavivirus; pestivirus; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 7..3678
XX /*tag= a
XX /product= "Breast cancer antigen Her2 variant"
XX
XX WO200240059-A2.
XX
XX 23-MAY-2002.
XX
XX

PF 01-NOV-2001; 2001WO-US45636.
XX
PR 01-NOV-2000; 2000US-0704232.
XX
XX (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.
PA (MINC/) MINCHEFF M S.
PA (LOUK/) LOUKINOV D I.
PA (ZOUB/) ZOUBAK S.
XX
XX Mincheff MS, Loukinov DI, Zoubak S;
PI WPI; 2002-527524/56.
XX P-PSDB; AAU98923.
DR
XX Inducing a cell-mediated immune response against a target antigen,
XX reducing undesired cells and stimulating presentation of an antigen by
XX a cell, comprises administering a polynucleotide encoding a variant of
XX an antigen -
XX
XX Disclosure; Page 128-134; 146pp; English.
XX
XX The invention relates to a method of inducing a cell-mediated immune
XX response against a cell comprising a target antigen (I) in a subject,
XX treating a subject having undesired cells, for example tumour cells
XX or virally infected cells (C), reducing the number of (C) in a subject,
XX and stimulating presentation of (I) by a cell. This is done by
XX administering a polynucleotide (II) encoding a variant of (I), so that
XX (II) expressed in a cell and cell-mediated immune response is induced.
XX The method can be used to treat prostate cancer, breast cancer,
XX colorectal cancer and pancreatic cancer, as well as lymphomas and
XX leukemias. The method is also useful in treating chronic viral
XX infections such as those caused by hepatitisviruses, lentiviruses
XX (including human immunodeficiency virus (HIV)), herpesviruses and the
XX flaviviruses and pestiviruses. The present sequence represents the coding
XX sequence of human breast cancer antigen, Her2 variant, used as a target
XX antigen in the method of the invention.
XX
XX Sequence 3678 BP; 746 A; 1137 C; 1091 G; 704 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 0.0581 Length: 3678
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 24 Gaps: 0
JUNC_SEQ8_SEQ5 (1-30) x ABR86207 (1-3678)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2869 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACACACCTTC 2922
RESULT 10
AAT40739
ID AAT40739 standard; cDNA; 3768 BP.
XX
XX AAT40739;
AC
XX 01-JAN-1997 (first entry)
XX
XX HER-2/neu oncogene.
DE
XX
XX HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
KW breast cancer; ovary cancer; colon cancer; lung cancer;
KW prostate cancer; genetic immunisation; tumour; vaccine; vector;
KW ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..3765
XX /*tag= b
XX
XX

```

FT      /note= "nucleotides 2026-3765 (claim 1) code for
FT      HER-2/neu intracellular domain"
PN      WO9630514-A1.
XX
XX      03-OCT-1996.
PD
XX
XX      28-MAR-1996; 96WO-US01689.
PF
XX
XX      31-MAR-1995; 95US-0414417.
PR
XX      (UNIW ) UNIV WASHINGTON.
PA
XX      Cheever MA, Disis ML;
XX      WPI; 1996-455361/45.
XX      P-PSDB; AAW01111.
DR
XX      DNA encoding HER-2-neu poly:peptide(s) - used for prevention or
PT      treatment of malignancies with which the HER-2/neu oncogene is
PT      associated
XX
XX      Claim 1; Page 49-56; 71pp; English.
PS
XX      Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or
CC      c-erbB2) protein (AAW01111). The oncogene is overexpressed in various
CC      cancers, including breast, ovarian, colon, lung and prostate, and
CC      appears to induce malignancies through quantitative mechanisms that
CC      result from increased or deregulated expression of an essentially
CC      normal gene product. Nucleotides 2026-3765 of the cDNA sequence
CC      code for the intracellular domain (lys676-Val1255) of the HER-2/neu
CC      protein, which is useful for immunisation against malignancy.
CC      Nucleic acids can be used to direct expression of the intracellular
CC      domain in transformed host cells, or are used, alone or in a viral
CC      vector, for genetic immunisation of an animal.
XX
XX      Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
SQ

```

```

FT      /*tag= b
FT      /note= "region which elicits immune response"
PN      US5869445-A.
XX
XX      09-FEB-1999.
PD
XX
XX      01-APR-1996; 96US-0625101.
PF
XX
XX      01-APR-1996; 96US-0625101.
PR
XX      17-MAR-1993; 93US-0033644.
PR
XX      12-AUG-1993; 93US-0106112.
PR
XX      31-MAR-1995; 95US-0414417.
XX
XX      (UNIW ) UNIV WASHINGTON.
XX
XX      Cheever MA, Disis ML;
XX      WPI; 1999-152835/13.
XX      P-PSDB; AAW92406.
DR
XX
XX      Use of HER-2/neu polypeptides - for eliciting an immune response to
PT      an HER-2/neu associated malignancy, particularly for treating or
PT      preventing tumours
XX
XX      Claim 1a; Column 23-32; 26pp; English.
PS
XX      This sequence encodes the human HER-2/neu oncogene protein. A fragment
CC      of this protein is used in a method for eliciting or enhancing an immune
CC      response to HER-2/neu protein. The polypeptide can stimulate T cells and
CC      B cells to produce an immune response to the HER-2/neu protein. The
CC      method can be used for immunisation against a malignancy in which the
CC      HER-2/neu oncogene is associated and in the treatment of an existing
CC      tumour, or to prevent tumour occurrence or reoccurrence.
XX
XX      Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
SQ

```

```

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 17 Gaps: 0

```

JUNC_SEQ8_SEQ5 (1-30) x AAT40739 (1-3768)

```

QY      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
      ||| |||||
Db      2959 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTTGGACAGCACCTTC 3012

```

```

RESULT 11
AAAX01912
ID      AAAX01912 standard; DNA; 3768 BP.
XX
XX      AAX01912;
AC
XX      21-APR-1999 (first entry)
DT
XX      Human HER-2/neu oncogene DNA.
DE
XX      HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
KW      malignancy; treatment; tumour; ss.
KW
XX      Homo sapiens.
OS
XX      Key Location/Qualifiers
FH      1..3768
FT      CDS
FT      /*tag= a
FT      /product= "HER-2/neu"
FT      /note= "oncogene"
FT      2026..3765
FT      misc_feature

```

```

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 20 Gaps: 0

```

JUNC_SEQ8_SEQ5 (1-30) x AAX01912 (1-3768)

```

QY      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
      ||| |||||
Db      2959 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTTGGACAGCACCTTC 3012

```

```

RESULT 12
AAAX09455
ID      AAAX09455 standard; DNA; 3768 BP.
XX
XX      AAX09455;
AC
XX      10-AUG-2000 (first entry)
DT
XX      Human heregulin 2 (Her2) coding sequence.
DE
XX      Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
KW      self-protein; cell-associated peptide antigen; foreign epitope;
KW      cancer; breast cancer; prostate cancer; ss.
KW
XX      Homo sapiens.
OS
XX      WO200020027-A2.
PN
XX      13-APR-2000.
PD
XX      05-OCT-1999; 99WO-DK00525.
XX

```

```
PR 05-OCT-1998; 98DK-0001261.
XX 20-OCT-1998; 98US-0105011.
XX (MEBI-) M & E BIOTECH AS.
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX WPI; 2000-349917/30.
DR P-PSDB; AAY93620.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
XX Claim 62; Page 187-193; 220pp; English.
XX
XX The claims detail a method for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, human
CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively.
XX
XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 21 Gaps: 0
JUNC_SEQ8_SEQ5 (1-30) x AAA09455 (1-3768)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012
RESULT 13
AAH23392
ID AAH23392 standard; DNA; 3768 BP.
XX
XX AAH23392;
XX
XX 25-SEP-2001 (first entry)
XX Human HER-2/neu protein encoding DNA.
XX
XX Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 1..3768
FT /*tag= a
FT /product= "HER-2/neu protein"
XX
XX WO200153463-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US01850.
```

```
XX 21-JAN-2000; 2000US-0177545.
XX (CORI-) CORIXA CORP.
XX Cheever MA, Hand-Zimmermann S;
PI WPI; 2001-476112/51.
XX P-PSDB; AAB85458.
XX
XX New antigen-presenting cells, useful as vaccines for eliciting or
PT enhancing an immune response to HER-2/neu protein, particularly useful
PT for treating or preventing cancer, e.g. breast cancer
XX
XX Claim 1; Page 41-46; 49pp; English.
XX
XX The invention provides an isolated antigen-presenting cell, which
CC expresses at least an immunogenic portion of a polypeptide that produces
CC an immune response to HER-2/neu protein. The antigen-presenting cells are
CC useful as vaccines for eliciting or enhancing an immune response to
CC HER-2/neu protein, particularly in treating or preventing malignancies in
CC which the HER-2/neu oncogene is associated. Specifically, these are
CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,
CC colon, lung or prostate cancers. The present sequence represents a DNA
CC encoding the human HER-2/neu protein (also known as p185 or c-erbB2).
XX
XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 22 Gaps: 0
JUNC_SEQ8_SEQ5 (1-30) x AAH23392 (1-3768)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012
RESULT 14
ABZ35744
ID ABZ35744 standard; DNA; 3768 BP.
XX
XX ABZ35744;
XX
XX 07-FEB-2003 (first entry)
XX Human ERBB2 polynucleotide SEQ ID NO 52.
XX
XX Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
KW protozoacide; gene expression; antisense; tumour; infection; plasmodium;
KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
KW Hepatitis C virus; human papilloma virus; gene; ds.
XX
XX Homo sapiens.
XX
XX DE10100568-A1.
XX
XX 18-JUL-2002.
XX
XX 09-JAN-2001; 2001DE-1000588.
XX
XX 09-JAN-2001; 2001DE-1000588.
XX (RIBO-) RIBOPHARMA AG.
XX
XX Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
XX WPI; 2002-683450/74.
XX
```

PT Inhibiting expression of target genes, useful e.g. for treating tumors,
PT by introducing into cells two double-stranded RNAs that are
PT complementary to the target
XX
XX Claim 13; Page 38-39; 100pp; German.
XX
CC The invention relates to inhibiting expression of a target gene in a cell
CC by introducing at least two oligoribonucleotides (dsRNAI and II), both
CC with a double-stranded (ds) structure of at most 49 sequential nucleotide
CC pairs. At least part of one strand (S1, S2) of the ds structures in each
CC of dsRNAI and II are complementary to regions in the target gene. The
CC method uses antisense inhibition of gene expression using double stranded
CC RNA inhibition (RNAi). The method is particularly used to treat tumours
CC or infections, especially by Plasmodium or viruses/viroids (pathogenic on
CC humans, animals or plants). The method provides more effective inhibition
CC of expression than known methods using a single dsRNA, even at very low
CC concentrations. When dsRNA has at least one unpaired nucleotide at the
CC end, stability (and thus effective concentration in the cell) is
CC improved and efficiency can be increased further by pretreating the cells
CC with interferon. The present sequence is that of a target DNA of the
CC invention.
XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 24 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x ABZ35744 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
DB 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 15
ABX09987
ID ABX09987 standard; DNA; 3768 BP.
XX
AC ABX09987;
XX
DT 23-JAN-2003 (first entry)
XX
DE Human ERBB2 DNA fragment SEQ ID 52.
XX
KW Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;
KW prion; inhibition; human; ds.
XX
OS Homo sapiens.
XX
PN DE10100587-C1.
XX
PD 21-NOV-2002.
XX
PF 09-JAN-2001; 2001DE-1000587.
XX
PR 09-JAN-2001; 2001DE-1000587.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
XX WPI; 2002-742209/81.
DR
XX
XX Inhibiting expression of target genes, e.g. oncogenes, in cells, by
PT introduction of complementary double-stranded oligoribonucleotide,
PT after treating the cell with interferon
XX
XX Disclosure; Page 43-44; 98pp; German.
PS
XX

CC This invention describes a novel method for inhibiting expression of a
CC target gene by introducing into the cell that contains the target gene
CC at least one oligoribonucleotide (dsRNAI) that has a double-stranded
CC (ds) structure of not more than 49 consecutive nucleotides (nt), where
CC at least a segment of one strand of the ds structure is complementary
CC with the target gene and the cells are treated with interferon before
CC introduction of dsRNAI. The method is used to inhibit expression of
CC target genes, particularly oncogenes, cytokine genes, Id (not defined)
CC protein genes; developmental or prion genes, or genes expressed in
CC pathogenic organisms (particularly plasmodia) or in viruses or viroids
CC (pathogenic in humans, animals or plants). Treating the cells with
CC interferon greatly increases the extent to which dsRNA can inhibit
CC expression of the target genes, and the effect is even greater when dsRNA
CC are modified to increase their stability. ABX09936-ABX10075 represent
XX gene fragments used to illustrate the method of the invention.

SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 24 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x ABX09987 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
DB 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGACCTTC 3012

Search completed: October 15, 2003, 23:48:52
Job time : 32.4941 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 432.026 Seconds
(without alignments)
2840.777 Million cell updates/sec

Title: JUNC_SEQ8_SEQ5

Perfect score: 160

Sequence: 1 ERGCPAEQRASPVTFQNEGLGPASPLDSTF 30

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Fgapop 10.0 , Fgapext 0.5

Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xlp

-MODEL-frame+p2n.model -DB-GenEmbl -Q/cgn2_1/USPTO_spool_p/HOLLERAN480/runat_15102003_131912_20336/app_query.fasta_1.4685
-UNITS-bits -START=1 -END=-1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480.ecgn_1_1_2724@runat_15102003_131912_20336 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.in.*
18: em.in.*
19: em.mu.*
20: em.or.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	147	91.9	2763	6	AX380942	AX380942 Sequence
2	147	91.9	2781	6	AX380944	AX380944 Sequence
3	83	51.9	1755	6	AX384609	AX384609 Sequence
4	83	51.9	1767	6	AX384607	AX384607 Sequence
5	83	51.9	1773	6	AX384610	AX384610 Sequence
6	83	51.9	1806	6	AX384608	AX384608 Sequence
7	83	51.9	3678	6	AX505114	AX505114 Sequence
8	83	51.9	3768	6	AR034479	AR034479 Sequence
9	83	51.9	3768	6	AR060704	AR060704 Sequence
10	83	51.9	3768	6	AX201817	AX201817 Sequence
11	83	51.9	3768	6	AX380923	AX380923 Sequence
12	83	51.9	3768	6	AX384604	AX384604 Sequence
13	83	51.9	3768	6	AX465456	AX465456 Sequence
14	83	51.9	3768	6	AX467229	AX467229 Sequence
15	83	51.9	3768	6	AX481438	AX481438 Sequence
16	83	51.9	3780	4	AB008451	AB008451 Canis fam
17	83	51.9	4473	6	AR080259	AR080259 Sequence
18	83	51.9	4473	6	AR167390	AR167390 Sequence
19	83	51.9	4473	9	HSEB2R	X03363 Human c-erb
20	83	51.9	4530	6	AR202597	AR202597 Sequence
21	83	51.9	4530	6	AR283481	AR283481 Sequence
22	83	51.9	4530	6	AX282577	AX282577 Sequence
23	83	51.9	4530	6	AX587649	AX587649 Sequence
24	83	51.9	4530	6	AX644071	AX644071 Sequence
25	83	51.9	4530	6	BD005474	BD005474 Cellular
26	83	51.9	4530	6	I21124	I21124 Sequence 9
27	83	51.9	4530	6	I59745	I59745 Sequence 9
28	83	51.9	4530	9	HUMHER2A	M11730 Human tyros
29	83	51.9	9274	6	AX060703	AX060703 Sequence
30	81	50.6	3771	6	AX189662	AX189662 Sequence
31	81	50.6	3771	6	AX380925	AX380925 Sequence
32	81	50.6	3955	6	AX380924	AX380924 Sequence
33	81	50.6	3955	6	I21129	I21129 Sequence 14
34	81	50.6	3955	6	I59750	I59750 Sequence 14
35	81	50.6	3955	10	RNNEUR	X03362 Rat mRNA fo
36	81	50.6	4694	10	BC046811	BC046811 Mus muscu
37	81	50.6	4695	10	BC053078	BC053078 Mus muscu
38	81	50.6	4727	10	AY116182	AY116182 Rattus no
39	80	50.0	4062	10	HAMNEU	D16295 Mesocricetu
40	79	49.4	13450	9	AB096612	AB096612 Homo sapi
41	79	49.4	16572	9	AB096614	AB096614 Homo sapi
42	79	49.4	20271	9	AB096613	AB096613 Homo sapi
43	79	49.4	30837	9	AY208911	AY208911 Homo sapi
44	79	49.4	161815	9	AC079199	AC079199 Homo sapi
45	79	49.4	168585	9	AC040933	AC040933 Homo sapi

ALIGNMENTS

RESULT 1

```

AX380942
LOCUS AX380942 2763 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 28 from Patent WO0212341.
ACCESSION AX380942
VERSION AX380942.1 GI:19575786
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
AUTHORS Cheever, M.A. and Gheysen, D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 28 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
Location/Qualifiers
1..2763
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD fusion protein cDNA"
BASE COUNT 571 a 855 c 772 g 565 t
ORIGIN
Alignment Scores:
Pred. No.: 3.46e-11 Length: 2763
Score: 147.00 Matches: 27
Percent Similarity: 96.67% Conservative: 2
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 91.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX380942 (1-2763)
Qy 1 GluArgGlyCysProAlaGluInArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACGAGAGCCAGCCAGTGTCTCAGAACGAGGACTTA 1977
Qy 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCCCTCCAGCCCTCCATGGACAGCACCTTC 2007
RESULT 2
AX380944
LOCUS AX380944 2781 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 30 from Patent WO0212341.
ACCESSION AX380944
VERSION AX380944.1 GI:19575787
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
AUTHORS Cheever, M.A. and Gheysen, D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 30 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
Location/Qualifiers
1..2781
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD-Tcp0 fusion protein cDNA"
BASE COUNT 574 a 859 c 779 g 569 t
ORIGIN
Alignment Scores:
Pred. No.: 3.49e-11 Length: 2781
Score: 147.00 Matches: 27
Percent Similarity: 96.67% Conservative: 2
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 91.88% Indels: 0
DB: 6 Gaps: 0

```

```

JUNC_SEQ8_SEQ5 (1-30) x AX380944 (1-2781)
Qy 1 GluArgGlyCysProAlaGluInArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACGAGAGCCAGCCAGTGTCTCAGAACGAGGACTTA 1977
Qy 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCCCTCCAGCCCTCCATGGACAGCACCTTC 2007
RESULT 3
AX384609
LOCUS AX384609 1755 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 6 from Patent WO0214503.
ACCESSION AX384609
VERSION AX384609.1 GI:19577810
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
TITLE Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 6 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..1755
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 376 a 517 c 531 g 331 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0303 Length: 1755
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX384609 (1-1755)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 937 GTGGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 990
RESULT 4
AX384607
LOCUS AX384607 1767 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 4 from Patent WO0214503.
ACCESSION AX384607
VERSION AX384607.1 GI:19577808
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
TITLE Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 4 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..1767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```


BASE COUNT 746 a 1137 c 1091 g 704 t
ORIGIN

FKGTPTAENPEYLGLDVPV"

Alignment Scores: 0.0666 Length: 3678
Pred. No.: 83.00 Matches: 16
Score: 83.00
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX505114 (1-3678)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2869 GTGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 2922

RESULT 8
AR034479
LOCUS AR034479 3768 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5869445.
ACCESSION AR034479
VERSION AR034479.1 GI:5950084
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3768)
AUTHORS Cheever, M.A. and Disis, M.L.
TITLE Methods for eliciting or enhancing reactivity to HER-2/neu protein
JOURNAL Patent: US 5869445-A 1 09-FEB-1999;
FEATURES Location/Qualifiers
1..3768
source /organism="unknown"

BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN

Alignment Scores: 0.0683 Length: 3768
Pred. No.: 83.00 Matches: 16
Score: 83.00
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AR034479 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 9
AX060704
LOCUS AX060704 3768 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 2 from Patent WO0100244.
ACCESSION AX060704
VERSION AX060704.1 GI:12406101
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Erickson, S. and Schwall, R.
TITLE Methods of treatment using anti-erbB antibody-maytansinoid
conjugates
JOURNAL Patent: WO 0100244-A 2 04-JAN-2001;
Genentech, Inc. (US)
FEATURES Location/Qualifiers
1..3768
source /organism="Homo sapiens"
/mol_type="genomic DNA"

BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN

/db_xref="taxon:9606"

Alignment Scores: 0.0683 Length: 3768
Pred. No.: 83.00 Matches: 16
Score: 83.00
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX060704 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 10
AX201817
LOCUS AX201817 3768 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0153463.
ACCESSION AX201817
VERSION AX201817.1 GI:15391666
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Cheever, M.A. and Hand-Zimmermann, S.
TITLE Compounds and methods for prevention and treatment of her-2/ neu
associated malignancies
JOURNAL Patent: WO 0153463-A 1 26-JUL-2001;
CORIAX CORPORATION (US)
FEATURES Location/Qualifiers
1..3768
source /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC60126.1"
/db_xref="GI:15391667"
/translation="MELAAALCRWGLLLALLPGAASTOYCTGTDMKLRPLASPETHLD
MRHLXGQGVQVQGNLELYLPYNASISFLDQIEVQGVYIAHNQVQVPLORLRIV
RGQLFEDNVALAVLDNGDPLNNTPTVTPGASGGLRELQLRSLTEILKGLVLRNPQ
LCYQDTILMKDIFHKNNQALALTIDNTRACHPSPCMKSGSCWSESDQSLRPT
VCAGGACRGKGLPTDCHEQCAAGCTGPKHSDCLACLFHNHSGICELCHCPALVTNT
DTFESMPNPEGRYTFGACSVTACPNYVLTVDGSCITLVCPLNHVEVTAEDGTORCEK
SKPCARVCYGLGMEHLREVRVTSANTQIEFAGCKIFGSLAFLPESFDGDPASNTAPL
QPEQLVFTLEITGYLISAWPDSLPDLSVFQNLQVIRGRILHNGAYSLTLQGLGI
SWIGLSRLRELGLALIHNTLCFVHTVPDQFNPHQALLTANRPEDECVEG
LACHQICARHGCPGPTQCVNCSQFLRGQECVECRVLQSLPREYVNAHCLPCHPE
CQPNQSVTCFCGPEADQVACAHYKDPFCVACRPSGVKPDLSYFWIKPDEEGACQ
PCPINCTHSCVDLDDKCPAEORASPLTISAVVGLLVVGLVFGVGLIKRQOKI
RKYTMRLLOETELVELTPSGAMPNAQMRILKTELKRVKVLGSGAGTVYKGIWI
PUGENVKIPVAKVLRNTSPKANKEILDEAYVMAGVSPVYSRLIGLICLTSTVQLVT
QLMPYGCILLDHVRENRLSGSLQDLNWCQIAKMSYLEDVRLVHLVLDLAARNLVKS
NHVITDFGLARLLDIDETEHADGKGVFKWMALESILRRFTHSDVMSYGVTVME
LMTFGAKPDGPAREIPDLLEKGERLPQPICTIDVYIMVWKWMDISRCYRFRREL
VSFESRMARDPQRFVIONEDLGASPLDSTFYRSLLEDDMDGLVADAEYLVPOQGF
FCPDPAAGAGGVHHRSSSTRSGGDLTLGLEPSEEAAPRSPAPSGAGSDVFDG
DUGMGAAGLQSLPHTDPSPLQRYSEDPTVLPSTEDGTGVAPLTCSPQPYVNPQDVR
POPPSPREGPLPAARPAATLERPKTLSPKNGVVKDVFAGGAVENPEYLTPOGGAA
POPHPPAFSPAFDNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVVPV"

1..3768
source

CDS

BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN

Alignment Scores: 0.0683 Length: 3768
Pred. No.: 83.00 Matches: 16
Score:

Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX201817 (1-3768)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCACCCAGTCCCTTGGACACACCTTC 3012

RESULT 11

AX380923
 LOCUS AX380923 3768 bp DNA linear PAT 18-MAR-2002
 DEFINITION Sequence 9 from Patent WO0212341.
 ACCESSION AX380923
 VERSION AX380923.1 GI:19575767
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Cheever, M.A. and Gheysen, D.
 Her-2/neu fusion proteins
 TITLE Patent: WO 0212341-A 9 14-FEB-2002;
 JOURNAL CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)

FEATURES

source
 1. .3768
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

CDS
 1. .3768
 /note="human Her-2/neu protein"
 /codon_start=1
 /protein_id="CAD28392.1"
 /db_xref="GI:19575768"

/translation="MELAALCRWGLLLALLPAGAASQVCTGDMKLRLLPASPEHLD
 MRHLVGGCQVQGNLELTPVNASLSFLQDQEVGYVLIHNRQVPLQRLRV
 LQVDTLWMDIFHKNQALTLIDTNRSRACHPCSPMCKGSCWGSSEDCSLRT
 VCAGGACRCGPIPTDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNT
 DTFSMPNGRYTFGASCTACPYNYLSDVGSCTLVCPHNOEVTAEHGTQCEKC
 SKPCARVYGLGMEHREVRVTSANTQEFAGCKIFGSLAFLPESDGPASNTAL
 QPGLQVFELEETGYLYISANPDSLPLDSVFQNLQVIRGRILHNGAYSLTQGLGI
 SWGLRLSELGSLALIHNTLHLCFVHTVPMQDLFRNPHOALLHTANRDECEVGG
 LACHLCARGHCGPGTQCVCNCSQFLRGQCEVEECVLRQGLPREYVNRHCLPCHPE
 COPNGSVTCFGEADQCVACAHYKDPFCVARGCPKPSVGLSYMPYTKPDEEGACQ
 PCPINCPSHCVDDDDKCPAERASPLTSIIISAVVGLLVVGLVFGILIKRQKI
 RKYTMRLLELTELVEPLTPSGAMPNOAMRILKTELKRVKVLGSGAFGVYKGIWI
 QLMYPGCLLDHVRNCRGLSQDLNWCMIKGMVLEDEVRVLRHDLAARNVLYKSP
 NHVKTDFGLARLLDIDETVHADGGKVPKMALESILRRRTHQSDVMSYGVYVWE
 LMTFGAKPIPAIVLRENTSPKANKEILDEAYVMAGVSPVSRLLGICLTSTVQLVT
 VSFSEMRAPQORFVYIQNEDLGPASPLDSTFYRSLLDEDDMDGLVDAEETLVPOQGF
 FCPDPAAGAGVHHRSSRSGGDLTLGLEPSEEPAPSLAPSEAGSDVDFG
 DLGMAKGLQSLPTHPDPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVR
 POPPSREGPLPAARPGATLERPKTLSPGKNGVVKDVFAGGAVENPEYLTTPQGGAA
 POPPPPAFPAFDNLYYWDPPPERGAPSTFKGTPTAENPEYLGIDVVPV"
 1. .1959

misc_feature

/note="extracellular domain (ECD) of human Her-2/neu"
 2026. .3765
 /note="intracellular domain (ICD) of human Her-2/neu"
 2968. .3765
 /note="phosphorylation domain (PD) of human Her-2/neu"
 2968. .3144
 /note="preferred portion of the phosphorylation domain
 (delta PD) of human Her-2/neu"

BASE COUNT 759 a 1171 c 1119 g 719 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.0683 Length: 3768

Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX380923 (1-3768)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCACCCAGTCCCTTGGACACACCTTC 3012

RESULT 12

AX384604
 LOCUS AX384604 3768 bp DNA linear PAT 19-MAR-2002
 DEFINITION Sequence 1 from Patent WO0214503.
 ACCESSION AX384604
 VERSION AX384604.1 GI:19577806
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
 McNeill, P.D. and Vedvick, T.S.
 Compositions and methods for the therapy and diagnosis of
 her-2/neu-associated malignancies
 Patent: WO 0214503-A 1 21-FEB-2002;
 JOURNAL CORIXA CORPORATION (US)

FEATURES

Location/Qualifiers
 1. .3768
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

CDS

1. .3768
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD28577.1"
 /translation="GI:19577807"
 /db_xref="GI:19577807"
 /translaton="MELAALCRWGLLLALLPAGAASQVCTGDMKLRLLPASPEHLD
 MRHLVGGCQVQGNLELTPVNASLSFLQDQEVGYVLIHNRQVPLQRLRV
 LQVDTLWMDIFHKNQALTLIDTNRSRACHPCSPMCKGSCWGSSEDCSLRT
 VCAGGACRCGPIPTDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNT
 DTFSMPNGRYTFGASCTACPYNYLSDVGSCTLVCPHNOEVTAEHGTQCEKC
 SKPCARVYGLGMEHREVRVTSANTQEFAGCKIFGSLAFLPESDGPASNTAL
 QPGLQVFELEETGYLYISANPDSLPLDSVFQNLQVIRGRILHNGAYSLTQGLGI
 SWGLRLSELGSLALIHNTLHLCFVHTVPMQDLFRNPHOALLHTANRDECEVGG
 LACHLCARGHCGPGTQCVCNCSQFLRGQCEVEECVLRQGLPREYVNRHCLPCHPE
 COPNGSVTCFGEADQCVACAHYKDPFCVARGCPKPSVGLSYMPYTKPDEEGACQ
 PCPINCPSHCVDDDDKCPAERASPLTSIIISAVVGLLVVGLVFGILIKRQKI
 RKYTMRLLELTELVEPLTPSGAMPNOAMRILKTELKRVKVLGSGAFGVYKGIWI
 QLMYPGCLLDHVRNCRGLSQDLNWCMIKGMVLEDEVRVLRHDLAARNVLYKSP
 NHVKTDFGLARLLDIDETVHADGGKVPKMALESILRRRTHQSDVMSYGVYVWE
 LMTFGAKPIPAIVLRENTSPKANKEILDEAYVMAGVSPVSRLLGICLTSTVQLVT
 VSFSEMRAPQORFVYIQNEDLGPASPLDSTFYRSLLDEDDMDGLVDAEETLVPOQGF
 FCPDPAAGAGVHHRSSRSGGDLTLGLEPSEEPAPSLAPSEAGSDVDFG
 DLGMAKGLQSLPTHPDPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVR
 POPPSREGPLPAARPGATLERPKTLSPGKNGVVKDVFAGGAVENPEYLTTPQGGAA
 POPPPPAFPAFDNLYYWDPPPERGAPSTFKGTPTAENPEYLGIDVVPV"
 759 a 1171 c 1119 g 719 t

BASE COUNT 759 a 1171 c 1119 g 719 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.0683 Length: 3768
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX384604 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012

RESULT 13

AX465456
LOCUS
DEFINITION
Sequence 1 from Patent WO0213847.
ACCESSION
AX465456
VERSION
AX465456.1 GI:21899819
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
Gaiger, A., Cheever, M.A. and Hand-zimmermann, S.
METHODS for diagnosis and therapy of hematological and
virus-associated malignancies
JOURNAL
Patent: WO 0213847-A.1 21-FEB-2002;
CORIXA CORPORATION (US)

FEATURES
source
1..3768
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD42363.1"
/db_xref="GI:21899820"
/translation="MELAAALCRWGLLLALLPPGAASCTVCTGTDMLRLPASPETHLD
MLRHLVQGCQVQGNLELYLPTNASLSFLQDIEQVGYVLIHNRVQVPLQRLRV
RGTOLEFEDNAYALVNDGDLPLNTTPTVGTASGGLRELQRLSLTELKGGVLQIRNPQ
LCYQDILMKDIFHKNNQLALTLIDNRSRACHPCSPCKGSRGSESSDCQSLRT
VCAGGACRKGPLPTDCHEQCAAGCTGPKSDCLACLFHNSGICELCHCPALVTYNT
DTPESNPGEGRYTFGACSVTACPYNYLSTDVSGCTLVCLPHNQEVTAEDGTQRCCK
SKPCARVCYGLGMEHUREVRAVTSANIQEFAGCKTFGLAFPLPESFDGDPASNTAPL
OPOLQVETLEETITGVLVISAWPDSLPLSVFQNLQVIRGRILHNGAYSLTQGLGI
SWGLSRLSRLSGSLALIHNTHLCEVHVVPMDLFRNPHQALLHTANRPECVGE
LACHQICARGHCGPCTQCVNCSQFLRQECVEECRVLQGLPREVNAHCLPCHPE
CQPNQSGTCFGEADQVACAHYKDPFCVAPSGVAPDLSYMPIMWKPDEEGACQ
PCPINTHSCVDLDDKCPAEQASPLTIIISAVVGLLVGLVVFGLIKRRQOKI
RKYTMRLLEQETELVPLTPSGAMPNQAMRILKETELRKVKVLSGAGFTVYKGLWI
PDGENVKIPVAIKVRENTSPKANKELIDEAAYMAGVGSFYVSRLLGICLTSTVQVLT
QLMPYCLLDHVRNRRGLSQDLNWCQIAKMSYLEDVRLVHRLDARNVLYKSP
NHVKITDFGLRLDIDETEHADGKVPKIKMALESILRRRTTHQSDVMSYGVTVME
LMTFGAKYDGPAREIPDLLEKGERLPQPICTIDYIMVIMVKWMDIDSECRPRREL
VSFESRMARDQRFVVIQNEDLGPASPLDSTFVRSLEDDMDGLVDAEYLVPOQGF
FCPDPAPGAGGMVHRRSSSTRSGGDLTLGLEPSEEPAPRSLAPSEAGSDVDFG
DLGMAAKGLQSLPTHDPSPLORYSEDPTVPLPSETDGYVAPLTCSQPQEVNQPDVR
PQPPSPREGPLPAARAGATLERAKTLSPKNGVVKDVFAGGAVENPEYLYTPQGGAA
POPHPPAPSPAFDNLVYWDQDPPPERGAPPSFTKGTPTAENPEYLGLDVVPV"

BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN

Alignment Scores:
Pred. No.: 0.0683 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX465456 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012

RESULT 14

AX467229

LOCUS
DEFINITION
Sequence 3 from Patent WO0234287.
ACCESSION
AX467229
VERSION
AX467229.1 GI:21900511
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
Beier, A.M., Gautam, A. and Mouritsen, S.R.
Novel therapeutic vaccine formulations
JOURNAL
Patent: WO 0234287-A 3 02-MAY-2002;
Pharmexa A/S (DK)

FEATURES
source
1..3768
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD42452.1"
/db_xref="GI:21900512"
/translation="MELAAALCRWGLLLALLPPGAASCTVCTGTDMLRLPASPETHLD
MLRHLVQGCQVQGNLELYLPTNASLSFLQDIEQVGYVLIHNRVQVPLQRLRV
RGTOLEFEDNAYALVNDGDLPLNTTPTVGTASGGLRELQRLSLTELKGGVLQIRNPQ
LCYQDILMKDIFHKNNQLALTLIDNRSRACHPCSPCKGSRGSESSDCQSLRT
VCAGGACRKGPLPTDCHEQCAAGCTGPKSDCLACLFHNSGICELCHCPALVTYNT
DTPESNPGEGRYTFGACSVTACPYNYLSTDVSGCTLVCLPHNQEVTAEDGTQRCCK
SKPCARVCYGLGMEHUREVRAVTSANIQEFAGCKTFGLAFPLPESFDGDPASNTAPL
OPOLQVETLEETITGVLVISAWPDSLPLSVFQNLQVIRGRILHNGAYSLTQGLGI
SWGLSRLSRLSGSLALIHNTHLCEVHVVPMDLFRNPHQALLHTANRPECVGE
LACHQICARGHCGPCTQCVNCSQFLRQECVEECRVLQGLPREVNAHCLPCHPE
CQPNQSGTCFGEADQVACAHYKDPFCVAPSGVAPDLSYMPIMWKPDEEGACQ
PCPINTHSCVDLDDKCPAEQASPLTIIISAVVGLLVGLVVFGLIKRRQOKI
RKYTMRLLEQETELVPLTPSGAMPNQAMRILKETELRKVKVLSGAGFTVYKGLWI
PDGENVKIPVAIKVRENTSPKANKELIDEAAYMAGVGSFYVSRLLGICLTSTVQVLT
QLMPYCLLDHVRNRRGLSQDLNWCQIAKMSYLEDVRLVHRLDARNVLYKSP
NHVKITDFGLRLDIDETEHADGKVPKIKMALESILRRRTTHQSDVMSYGVTVME
LMTFGAKYDGPAREIPDLLEKGERLPQPICTIDYIMVIMVKWMDIDSECRPRREL
VSFESRMARDQRFVVIQNEDLGPASPLDSTFVRSLEDDMDGLVDAEYLVPOQGF
FCPDPAPGAGGMVHRRSSSTRSGGDLTLGLEPSEEPAPRSLAPSEAGSDVDFG
DLGMAAKGLQSLPTHDPSPLORYSEDPTVPLPSETDGYVAPLTCSQPQEVNQPDVR
PQPPSPREGPLPAARAGATLERAKTLSPKNGVVKDVFAGGAVENPEYLYTPQGGAA
POPHPPAPSPAFDNLVYWDQDPPPERGAPPSFTKGTPTAENPEYLGLDVVPV"

BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN

Alignment Scores:
Pred. No.: 0.0683 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX467229 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012

RESULT 15

AX481438
LOCUS
DEFINITION
Sequence 52 from Patent WO02055693.
ACCESSION
AX481438
VERSION
AX481438.1 GI:22316352
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

AX481438
LOCUS
DEFINITION
Sequence 52 from Patent WO02055693.
ACCESSION
AX481438
VERSION
AX481438.1 GI:22316352
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.
Method for inhibiting the expression of a target gene
Patent: WO 02055693-A 52 18-JUL-2002;

JOURNAL

Ribopharma AG (DE)

FEATURES

Location/Qualifiers

1..3768

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 758 a 1170 c 1121 g 719 t

ORIGIN

Alignment Scores:
Pred. No.: 0.0683 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX481438 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 2959 GTGGTCATCCAGAAATGAGGACTTGGCCCGAGCCAGTCCCTTGGACAGCACCTTC 3012

Search completed: October 16, 2003, 11:01:50
Job time : 436.208 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 234.794 Seconds
(without alignments)
3105.423 Million cell updates/sec

Title: JUNC_SEQ8_SEQ4
Perfect score: 160
Sequence: 1 ERGCPAEQRASPVTFQNEGLGPASPLDSTFF 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-xlp
-Q/cgn2_1/USPTO_spool_p/HOLLERAN480/runat_15102003_131913_20548/app_query.fasta_1.4685
-DB-EST -QFMT-fastap -SUFFIX-rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cgi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER-HOLLERAN480 -CGN_1_112645 -runat_15102003_131913_20548 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estimu.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hct.*
9: gb_est1.*
10: gb_est2.*
11: gb_hct.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_nam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gssl.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	83	51.9	278	10	BF923905	BF923905 QV4-CT025
2	83	51.9	294	10	BF757857	BF757857 CM4-CT057
3	83	51.9	337	10	BE091682	BE091682 IL2-BT073
4	83	51.9	348	9	A1909847	A1909847 QV-BT225-
5	83	51.9	378	14	CB266492	CB266492 1005398 H
6	83	51.9	423	12	BM703962	BM703962 UI-E-CK1-
7	83	51.9	464	2	HSM067221	Bx477782 Homo sapi
8	83	51.9	496	12	BM795191	BM795191 K-EST0076
9	83	51.9	505	12	BM854421	BM854421 K-EST0136
10	83	51.9	531	12	BM830170	BM830170 K-EST0103
11	83	51.9	536	14	CB129403	CB129403 K-EST0179
12	83	51.9	547	12	BM787824	BM787824 K-EST0066
13	83	51.9	567	2	HSM077766	Bx487579 Homo sapi
14	83	51.9	574	12	BM829991	BM829991 K-EST0103
15	83	51.9	630	2	HSM073298	Bx483130 Homo sapi
16	83	51.9	691	14	CB853376	CB853376 UI-CF-FNO
17	83	51.9	767	10	BF240297	BF240297 601905830
18	83	51.9	849	14	CA489799	CA489799 AGENCOURT
19	83	51.9	852	14	CD516283	CD516283 AGENCOURT
20	83	51.9	866	14	CA488343	CA488343 AGENCOURT
21	83	51.9	893	14	CA455141	CA455141 AGENCOURT
22	83	51.9	894	14	CA454131	CA454131 AGENCOURT
23	83	51.9	899	14	CD558386	CD558386 AGENCOURT
24	83	51.9	916	13	B0845369	B0845369 AGENCOURT
25	83	51.9	919	13	B0845391	B0845391 AGENCOURT
26	83	51.9	924	14	CA489100	CA489100 AGENCOURT
27	83	51.9	932	14	CA487981	CA487981 AGENCOURT
28	83	51.9	984	14	CA489084	CA489084 AGENCOURT
29	83	51.9	1004	14	CD515356	CD515356 AGENCOURT
30	83	51.9	4715	11	AF318349	AF318349 Homo sapi
31	81	50.6	220	9	AW763060	AW763060 ur60c11.y
32	81	50.6	377	9	AW415355	AW415355 49593 MAR
33	81	50.6	483	4	BX530001	BX530001 RZPD Mus
34	81	50.6	518	10	BF039952	BF039952 Bp250015A
35	81	50.6	552	10	BF022141	BF022141 uy47c11.y
36	81	50.6	654	12	B1648646	B1648646 603277149
37	81	50.6	655	12	B1557778	B1557778 603236957
38	81	50.6	695	12	B1557170	B1557170 603239144
39	81	50.6	718	12	B1108229	B1108229 602902263
40	81	50.6	749	12	B1554781	B1554781 603236758
41	81	50.6	754	12	B1651822	B1651822 603299821
42	81	50.6	771	12	B1154609	B1154609 602904778
43	81	50.6	891	10	BF101641	BF101641 601753564
44	81	50.6	3110	11	AK031542	AK031542 Mus muscu
45	81	50.6	3372	11	BC046553	BC046553 Mus muscu

ALIGNMENTS

RESULT 1
BF923905
LOCUS QV4-NT0251-251100-599-all NF0251 Homo sapiens cdna, mRNA sequence.
DEFINITION BF923905
ACCESSION BF923905
VERSION BF923905.1 GI:12319793
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 278)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

PUBLISHED

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV4&t2=QV4-NT0251-

251100-599-all&t3=2000-11-25&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 10

High quality sequence stop: 278.

Location/Qualifiers

1. .278

/organism="Homo sapiens"

/mol_type="mrna"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="NT0251"

/note="Organ: nervous_tumor; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT 52 a 78 c 79 g 69 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0896 Length: 278

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 10 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x BF923905 (1-278)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 135 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 188

RESULT 2

BF757857

LOCUS

CM4-CT0574-101100-428-f11 CT0574 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BF757857

VERSION

BF757857.1 GI:12105757

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 294)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

PUBLISHED

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&t2=CM4-CT0574-

101100-428-f11&t3=2000-11-10&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 9

High quality sequence stop: 294.

Location/Qualifiers

1. .294

/organism="Homo sapiens"

/mol_type="mrna"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="CT0574"

/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:

SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters patent application No. 196

,716 - Ludwig Institute for Cancer Research) profiles

into the puc 18 vector. Reverse transcription of tissue

mRNA and cDNA amplification were performed under low

stringency conditions."

BASE COUNT 55 a 85 c 86 g 68 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0963 Length: 294

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 10 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x BF757857 (1-294)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 129 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 182

RESULT 3

BE091682/c

LOCUS

IL2-BT0731-240400-070-E05 BT0731 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BE091682

VERSION

BE091682.1 GI:8482134

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 337)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE
10737800
PUBMED
COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ft2=IL2-BT0731-240
400-070-E05ft3=2000-04-24&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 336.

FEATURES
source

1..337
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BT0731"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 72 a 101 c 96 g 68 t
ORIGIN

Alignment Scores:
Pred. No.: 0.115 Length: 337
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x BE091682 (1-337)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 148 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTC 95

RESULT 4
AI909847

LOCUS QV-BT225-050599-038 BT225 Homo sapiens cDNA, mRNA linear EST 30-MAR-2000
DEFINITION QV-BT225-050599-038 BT225 Homo sapiens cDNA, mRNA sequence.
ACCESSION AI909847
VERSION AI909847.1 GI:6500527
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/seq/gethtml.pl?tl=QV&t2=QV-BT225-038.html
&t3=050599&t4=1)
Seq primer: puc 18 forward.

FEATURES
source

1..348
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="female"
/dev_stage="Adult"
/clone_lib="BT225"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 68 a 96 c 114 g 70 t
ORIGIN

Alignment Scores:
Pred. No.: 0.12 Length: 348
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 9 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AI909847 (1-348)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 15 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTC 68

RESULT 5
CB266492

LOCUS 1003398 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
DEFINITION CDNA 5', mRNA sequence.
ACCESSION CB266492
VERSION CB266492.1 GI:28441078
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Yang,R.-Z., Shuldiner,A. and Gong,D.-W.
1 (bases 1 to 378)
TITLE EST analysis of human adipose gene expression
JOURNAL Unpublished
COMMENT Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition
University of Maryland
660 Redwood St. HH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgong@medicine.umaryland.edu
PCR PRimers
FORWARD: CTCGGGAAGCGCCCATCTGTTGTG
BACKWARD: AATAGCACTCACTATAGGCGAATTG
Seq primer: GTTGGTACCCGGGAATTC.

FEATURES
source

1..378
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"


```

/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/notes="vector: lambdaTriplex"
BASE COUNT 71 a 110 c 117 g 80 t
ORIGIN
Alignment Scores:
Pred. No.: 0.133 Length: 378
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
Gaps: 0
JUNC_SEQ8_SEQ4 (1-30) x CB266492 (1-378)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 124 GTGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTGGACAGCACCTTC 177
RESULT 6
LOCUS
DEFINITION
BM703962 423 bp mRNA linear EST 28-FEB-2002
UI-E-CK1-afk-o-11-0-UI.r1 UI-E-CK1 Homo sapiens cDNA clone
BM703962
VERSION
BM703962.1 GI:19017220
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 423)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel.: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1. .423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CK1-afk-o-11-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK1"
/notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site:1: EcoR I; Site:2: Not I;
UI-E-CK1 is a normalized cDNA library containing the
following tissue(s): Retina Foveal and Macular. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated

```

to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTCC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 78 a 129 c 131 g 85 t
ORIGIN

Alignment Scores:
Pred. No.: 0.154 Length: 423
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x BM703962 (1-423)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 147 GTGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTGGACAGCACCTTC 200

RESULT 7

ID HSM067221 standard; RNA; EST; 464 BP.

XX AC BX477782;

XX SV BX477782.1

XX DT 09-MAY-2003 (Rel. 75, Created)

XX DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

XX DE Homo sapiens mRNA; EST DKFZp686M12198_r1 (from clone DKFZp686M12198)

XX KW EST; expressed sequence tag.

XX OS Homo sapiens (human)

XX OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia;

XX OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX RN [1]

XX RA Bioecker H., Boecker M., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;

XX RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.

XX RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

XX CC This is the 5' sequence of the clone insert

CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

CC sequenced by GBF (National Research Centre for Biotechnology

CC Ltd., Braunschweig/Germany) within the cDNA sequencing

CC consortium of the German Genome Project.

CC No sl sequence available.

CC This clone (DKFZp686M12198) is available at the RZPD in Berlin.

CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,

CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX Key Location/Qualifiers

FT source 1. .464

FT /db_xref="taxon:9606"

FT /mol_type="mRNA"

FT /organism="Homo sapiens"

FT /clone="DKFZp686M12198"

FT /clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host

FT DH10B; sites sfiIA + sfiIB"

FT /dev_stage="adult"

FT /tissue_type="cdna-collection"
 XX Sequence 464 BP; 87 A; 127 C; 149 G; 101 T; 0 other;

Alignment Scores:
 Pred. No.: 0.173 Length: 464
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 Db: 2 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x HSM067221 (1-464)

OY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| |||||
 Db 248 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 301

RESULT 8
 BM795191
 LOCUS BM795191 496 bp mRNA linear EST 05-MAR-2002
 DEFINITION K-EST0076753 S21SNU520 Homo sapiens cDNA clone S21SNU520-26-D05 5',
 mRNA sequence.
 ACCESSION BM795191
 VERSION BM795191.1 GI:19143423
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 496)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 26 row: D column: 05
 High quality sequence stop: 496.
 Location/Qualifiers
 1. .496

FEATURES
 source
 1. .496
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S21SNU520-26-D05"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="Floating aggregates"
 /cell_line="SNU-520"
 /lab_host="Top10F"
 /clone_lib="S21SNU520"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
 Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tabacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

BASE COUNT 94 a 134 c 159 g 109 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.189 Length: 496
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 Db: 12 Gaps: 0
 JUNC_SEQ8_SEQ4 (1-30) x BM795191 (1-496)

OY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| |||||
 Db 334 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 387

RESULT 9
 BM854421
 LOCUS BM854421 505 bp mRNA linear EST 06-MAR-2002
 DEFINITION K-EST0136797 S21SNU520 Homo sapiens cDNA clone S21SNU520-57-G11 5',
 mRNA sequence.
 ACCESSION BM854421
 VERSION BM854421.1 GI:19210820
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 505)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 57 row: G column: 11
 High quality sequence stop: 505.
 Location/Qualifiers
 1. .505
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S21SNU520-57-G11"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="Floating aggregates"
 /cell_line="SNU-520"
 /lab_host="Top10F"
 /clone_lib="S21SNU520"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
 Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tabacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

BASE COUNT 95 a 138 c 161 g 111 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.193 Length: 505
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x BM854421 (1-505)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 Db 334 GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGTCCCTTGGACAGCACCTTC 387
 RESULT 10
 BM830170
 LOCUS
 DEFINITION K-EST0103592 S21SN520s1 531 bp mRNA linear EST 06-MAR-2002
 5', mRNA sequence.
 ACCESSION BM830170
 VERSION BM830170.1 GI:19186579
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 531)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 19 row: D column: 03
 High quality sequence stop: 531.
 Location/Qualifiers
 1..531
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S21SN520s1-19-D03"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="Floating aggregates"
 /cell_line="SNU-520"
 /lab_host="Top10F"
 /clone_lib="S21SN520s1"
 /notes="Organ: Stomach; Vector: pTZ18RP1; Site.1: EcoRI;
 Site.2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library. After analyzing and
 sequencing about 2,000 ~ 3,000 colonies in original cDNA
 library, the abundant cDNAs were selected and amplified by

PCR reaction using vector region primer including T7
 promoter as 5' primer and N(GT)14 as 3' primer. The PCR
 products were used as template for synthesis of
 biotinylated single stranded RNA by in vitro transcription
 reaction. The synthesized RNA probes were hybridized with
 antisense single stranded cDNAs prepared from original
 library and incubated with avidin-gel. After removing
 DNA-RNA hybrids by centrifuge, the subtracted cDNA
 libraries were constructed by transfection of the
 remaining DNA into competent cells E. coli Top10F' with
 electroporation method."

BASE COUNT 102 a 149 c 166 g 114 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0.206 Length: 531
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x BM830170 (1-531)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 Db 334 GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGTCCCTTGGACAGCACCTTC 387
 RESULT 11
 CB129403
 LOCUS
 DEFINITION K-EST0179037 C1SNU17 Homo sapiens cDNA clone C1SNU17-5-C08 5', mRNA
 sequence.
 ACCESSION CB129403
 VERSION CB129403.1 GI:28093000
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 536)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 5 row: C column: 08
 High quality sequence stop: 536.
 Location/Qualifiers
 1..536
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="C1SNU17-5-C08"
 /sex="F"
 /tissue_type="uterine"
 /cell_type="Epithelial"
 /cell_line="SNU-17"
 /lab_host="Top10F"
 /clone_lib="C1SNU17"
 /notes="Organ: Cervix; Vector: pCNS-D2; Site.1: EcoRI;
 Site.2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including
 EcoRI site by treatment of T4 RNA ligase and the first
 strand cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library. After analyzing and
 sequencing about 2,000 ~ 3,000 colonies in original cDNA
 library, the abundant cDNAs were selected and amplified by

FEATURES
 sourceFEATURES
 source

strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 105 a 150 c 167 g 114 t

ORIGIN

Alignment Scores:

Pred. No.: 0.209 Length: 536
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 14 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x CB129403 (1-536)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 339 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTTGGCAGACACCTTC 392

RESULT 12

BM787824

LOCUS K-EST0066898 S11SNUI Homo sapiens cDNA clone S11SNUI-23-G01 5',
DEFINITION mRNA sequence.

ACCESSION BM787824.1 GI:19136056

VERSION EST.

KEYWORDS SOURCE

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 547)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished

COMMENT Contact: Kim YS

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 23 row: G column: 01

High quality sequence stop: 547.

Location/Qualifiers

1..547

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S11SNUI-23-G01"

/sex="M"

/tissue_type="Stomach"

/cell_type="Lymphoblast-like"

/cell_line="SNU-1"

/lab_host="Top10F"

/clone_lib="S11SNUI"

/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI

oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18S-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 102 a 187 c 163 g 95 t

ORIGIN

Alignment Scores:

Pred. No.: 0.214 Length: 547
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x BM787824 (1-547)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 39 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTTGGCAGACACCTTC 92

RESULT 13

HSN077766

ID HSN077766 standard; RNA; EST; 567 BP.

XX AC BX487579;

SV BX487579.1

XX

DT 09-MAY-2003 (Rel. 75, Created)

DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

XX

DE Homo sapiens mRNA; EST DKFZp686M16257_r1 (from clone DKFZp686M16257)
XX EST; expressed sequence tag.

KW

XX Homo sapiens (human)

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX

RN

RP 1-567

RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,

RT Han M., Wiemann S.;

RL Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.

RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

XX

CC This is the 5' sequence of the clone insert
CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC sequenced by BMFZ (Biomedical Research Center at the
CC Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
CC sequencing consortium of the German genome Project.
CC No si sequence available.

CC This clone (DKFZp686M16257) is available at the RZPD in Berlin.
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX

Key Location/Qualifiers

FT source

1..567

/db_xref="taxon:9606"

/mol_type="mRNA"

/organism="Homo sapiens"

/clone="DKFZp686M16257"

FT /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
FT DH10B; sites SfiIA + SfiIB"
FT /dev_stage="adult"

FT /tissue_type="cDNA-collection"

XX Sequence 567 BP; 115 A; 146 C; 184 G; 122 T; 0 other;

Alignment Scores:

Pred. No.: 0.224 Length: 567
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 2 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x HSM077766 (1-567)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 Db 468 GTGGTCATCCAGATGAGGACTGGGCCCGCCAGTCCCTGGACACACCTTC 521

RESULT 14

BM829991

LOCUS BM829991 574 bp mRNA linear EST 06-MAR-2002
 DEFINITION K-EST0103390 S21SNU520s1 Homo sapiens cDNA clone S21SNU520s1-17-C05
 5', mRNA sequence.

ACCESSION BM829991

VERSION BM829991.1 GI:19186400

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,

Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and

Kim, Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished

COMMENT Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 17 row: C column: 05

High quality sequence stop: 574.

FEATURES

source

1..574
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S21SNU520s1-17-C05"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="Floating aggregates"
 /cell_line="SNU-520"
 /lab_host="Top10f"
 /clone_lib="S21SNU520s1"
 /note="Origin: Stomach; Vector: pTZ19Rpl; Site:1: EcoRI;
 Site:2: NotI. The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tabacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 converted cDNA vectors were used for transformation of
 competent cells E. coli Top10f' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library. After analyzing and

sequencing about 2,000 ~ 3,000 colonies in original cDNA
 library, the abundant cDNAs were selected and amplified by
 PCR reaction using vector region primer including T7 PCR
 promotor as 5' primer and N(dt)14 as 3' primer. The PCR
 products were used as template for synthesis of
 biotinylated single stranded RNA by in vitro transcription
 reaction. The synthesized RNA probes were hybridized with
 antisense single stranded cDNAs prepared from original
 library and incubated with avidin-gel. After removing
 DNA-RNA hybrids by centrifuge, the subtracted cDNA
 libraries were constructed by transformation of the
 remaining DNA into competent cells E. coli Top10f' with
 electroporation method."

BASE COUNT 109 a 159 c 185 g 121 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.228 Length: 574
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x BM829991 (1-574)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 Db 334 GTGGTCATCCAGATGAGGACTGGGCCCGCCAGTCCCTGGACACACCTTC 387

RESULT 15

HSM073298

ID HSM073298 standard; RNA; EST; 630 BP.

XX BX483130;

XX BX483130.1

XX 09-MAY-2003 (Rel. 75, Created)

DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

DE Homo sapiens mRNA; EST DKFZp686N09234_r1 (from clone DKFZp686N09234)

XX EST; expressed sequence tag.

XX Homo sapiens (human)

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX [1]

RP 1-630

RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Mewes H.W., Weil B.,

RA Anld C., Osanger A., Fobo G., Han M., Wiemann S.;

RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.

RL MIPS, Ingolstaeter Landstr.1, D-85764 Neuherberg, GERMANY

XX This is the 5' sequence of the clone insert

CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

CC Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;

CC sequenced by Medigenomix (Martinsried/Germany) within the cDNA

CC sequencing consortium of the German Genome Project.

CC No sl sequence available.

CC This clone (DKFZp686N09234) is available at the RZPD in Berlin.

CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,

CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX Key Location/Qualifiers

PH 1..630

FT source

FT /db_xref="taxon:9606"

FT /mol_type="mRNA"

FT /organism="Homo sapiens"

FT /clone="DKFzp686N09234"
FT /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
FT DH108; sites SfiIA + SfiIB"
FT /dev_stage="adult"
FT /tissue_type="cDNA-collection"
XX

SQ Sequence 630 BP; 124 A; 174 C; 196 G; 133 T; 3 other;

Alignment Scores:

Pred. No.:	0.257	Length:	630
Score:	83.00	Matches:	16
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	2
Query Match:	51.88%	Indels:	0
DB:	2	Gaps:	0

JUNC_SEQ8_SEQ4 (1-30) x HSM073298 (1-630)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
|||
Db 439 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTC 492

Search completed: October 16, 2003, 17:04:06
Job time : 236.794 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 11:08:07 ; Search time 31.0111 Seconds
(without alignments)
2540.503 Million cell updates/sec

Title: JUNC_SEQ8_SEQ4

Perfect score: 160

Sequence: 1 ERGCPAQRASPVTFQNEGLGPASPLDSTF 30

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/HOLLERAN480/runat_15102003_131915_20662/app_query.fasta_1.4685
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NAXLEN=2000000000 -USER=HOLLERAN480 -CGN_1_1_397 @runat_15102003_131915_20662
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PublishedApplications_NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	83	51.9	1115	14	US-10-102-806-165 Sequence 165, App

2	83	51.9	1713	12	US-10-378-393-14 Sequence 14, Appl
3	83	51.9	1755	10	US-09-930-125-6 Sequence 6, Appli
4	83	51.9	1767	10	US-09-930-125-4 Sequence 4, Appli
5	83	51.9	1773	10	US-09-930-125-7 Sequence 7, Appli
6	83	51.9	1806	10	US-09-930-125-5 Sequence 5, Appli
7	83	51.9	2411	12	US-10-378-393-10 Sequence 10, Appli
8	83	51.9	3765	12	US-10-207-498-5 Sequence 5, Appli
9	83	51.9	3768	9	US-09-811-123-8 Sequence 8, Appli
10	83	51.9	3768	9	US-09-811-115-2 Sequence 2, Appli
11	83	51.9	3768	10	US-09-854-356-9 Sequence 9, Appli
12	83	51.9	3768	10	US-09-930-125-1 Sequence 1, Appli
13	83	51.9	3768	12	US-10-313-644-1 Sequence 1, Appli
14	83	51.9	4473	11	US-09-441-411-5 Sequence 5, Appli
15	83	51.9	4473	12	US-10-101-510-81 Sequence 81, Appli
16	83	51.9	4473	14	US-10-146-473-32 Sequence 32, Appli
17	83	51.9	4473	14	US-10-207-655-44 Sequence 44, Appli
18	83	51.9	4530	10	US-09-877-177-11 Sequence 11, Appli
19	83	51.9	4530	12	US-10-007-926A-119 Sequence 119, App
20	83	51.9	4530	12	US-10-101-510-124 Sequence 124, App
21	83	51.9	4530	12	US-10-338-730-1 Sequence 1, Appli
22	83	51.9	4530	14	US-10-177-293-125 Sequence 125, App
23	83	51.9	4543	10	US-09-769-508-1 Sequence 1, Appli
24	83	51.9	4606	12	US-09-971-392-70 Sequence 70, Appli
25	83	51.9	4642	14	US-10-198-846-10896 Sequence 10896, A
26	83	51.9	9274	9	US-09-811-123-7 Sequence 7, Appli
27	83	51.9	9274	9	US-09-811-115-1 Sequence 1, Appli
28	81	50.6	3771	10	US-09-854-356-11 Sequence 11, Appli
29	81	50.6	3955	10	US-09-870-759-117 Sequence 117, App
30	81	50.6	3955	10	US-09-854-356-10 Sequence 10, Appli
31	81	50.6	3955	12	US-09-751-708A-117 Sequence 117, App
32	66	41.2	201	13	US-10-109-213-3 Sequence 3, Appli
33	58.5	36.6	143601	12	US-09-855-824-3 Sequence 3, Appli
34	58	36.2	1938	12	US-10-422-264-25 Sequence 25, Appli
35	58	36.2	2316	12	US-10-422-264-27 Sequence 27, Appli
36	58	36.2	2604	12	US-10-422-264-23 Sequence 23, Appli
37	58	36.2	3320	10	US-09-862-658-1 Sequence 1, Appli
38	58	36.2	3320	14	US-10-175-696-22 Sequence 22, Appli
39	58	36.2	3384	12	US-10-422-264-29 Sequence 29, Appli
40	57.5	35.9	363	11	US-09-918-995-28706 Sequence 28706, A
c 41	57	35.6	23432	9	US-09-764-869-1332 Sequence 1332, Ap
c 42	57	35.6	23432	14	US-10-091-504-1332 Sequence 1332, Ap
c 43	56.5	35.3	1557	10	US-09-774-434-6 Sequence 6, Appli
c 44	56.5	35.3	22452	10	US-09-764-868-1487 Sequence 1487, Ap
c 45	56.5	35.3	22452	10	US-09-764-868-1489 Sequence 1489, Ap

ALIGNMENTS

RESULT 1
US-10-102-806-165
; Sequence 165, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 165
; LENGTH: 1115
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (390)

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-165

Alignment Scores:
Pred. No.: 0.00141 Length: 1115
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 14 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-10-102-806-165 (1-1115)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 446 GTGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 499

RESULT 2

US-10-378-393-14
; Sequence 14, Application US/10378393
; Publication No. US20030182668A1
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-14

Alignment Scores:
Pred. No.: 0.00225 Length: 1713
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-10-378-393-14 (1-1713)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 904 GTGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 957

RESULT 3

US-09-930-125-6
; Sequence 6, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; OF HER-2/NEU-ASSOCIATED MALIGNANCIES

; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930.125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-6

Alignment Scores:
Pred. No.: 0.00231 Length: 1755
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-930-125-6 (1-1755)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 937 GTGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 990

RESULT 4

US-09-930-125-4
; Sequence 4, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930.125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-4

Alignment Scores:
Pred. No.: 0.00233 Length: 1767
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-930-125-4 (1-1767)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 937 GTGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 990

RESULT 5

US-09-930-125-7
; Sequence 7, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.

```

; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-7

Alignment Scores:
Pred. No.: 0.00233 Length: 1773
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-930-125-7 (1-1773)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 961 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCCACTCCCTTGGACAGCACCTTC 1014

RESULT 6
US-09-930-125-5
; Sequence 5, Application US/099301025
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-5

Alignment Scores:
Pred. No.: 0.00238 Length: 1806
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-930-125-5 (1-1806)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 994 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCCACTCCCTTGGACAGCACCTTC 1047

RESULT 7
US-10-378-393-10
; Sequence 10, Application US/10378393
; Publication No. US20030182668A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TYROSINIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVAT
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-10

Alignment Scores:
Pred. No.: 0.00327 Length: 2411
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-10-378-393-10 (1-2411)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 1602 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCCACTCCCTTGGACAGCACCTTC 1655

RESULT 8
US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; FILE REFERENCE: 30448.103-US-U1
; CURRENT APPLICATION NUMBER: US/10/207,498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5

Alignment Scores:
Pred. No.: 0.00532 Length: 3765
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-10-207-498-5 (1-3765)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

```

```
Db      2959 GTGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012
|||||
RESULT 9
US-09-811-123-8
; Sequence 8, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; TITLE OF INVENTION: ANTIBODY-WAYTANSINOID CONJUGATES
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811.123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-123-8

Alignment Scores:
Pred. No.:      0.00532      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    51.88%      Indels:      0
DB:             9          Gaps:          0

JUNC_SEQ08_SEQ4 (1-30) x US-09-811-123-8 (1-3768)

Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db      2959 GTGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012
|||||
RESULT 10
US-09-811-115-2
; Sequence 2, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811.115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred. No.:      0.00532      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    51.88%      Indels:      0
DB:             9          Gaps:          0

JUNC_SEQ08_SEQ4 (1-30) x US-09-811-115-2 (1-3768)

Db      2959 GTGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012
|||||
RESULT 11
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.:      0.00532      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    51.88%      Indels:      0
DB:             10         Gaps:          0

JUNC_SEQ08_SEQ4 (1-30) x US-09-854-356-9 (1-3768)

Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db      2959 GTGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012
|||||
RESULT 12
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
```

```
Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db      2959 GTGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012
|||||
RESULT 11
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.:      0.00532      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    51.88%      Indels:      0
DB:             10         Gaps:          0

JUNC_SEQ08_SEQ4 (1-30) x US-09-854-356-9 (1-3768)

Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db      2959 GTGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012
|||||
RESULT 12
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
 ; FILE REFERENCE: 210121.544
 ; CURRENT APPLICATION NUMBER: US/09/930,125
 ; CURRENT FILING DATE: 2001-08-14
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 3768
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(3765)
 US-09-930-125-1

Alignment Scores:
 Pred. No.: 0.00532 Length: 3768
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 10 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-930-125-1 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| |||||
 Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 3012

RESULT 13

US-10-313-644-1
 ; Sequence 1, Application US/10313644
 ; Publication No. US20030157119A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Cheever, Martin A.
 ; APPLICANT: Hand-Zimmerman, Susan
 ; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
 ; TITLE OF INVENTION: AND VIRUS-ASSOCIATED MALIGNANCIES
 ; FILE REFERENCE: 210121.483C3
 ; CURRENT APPLICATION NUMBER: US/10/313,644
 ; CURRENT FILING DATE: 2002-12-04
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 3768
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(3765)
 US-10-313-644-1

Alignment Scores:
 Pred. No.: 0.00532 Length: 3768
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-10-313-644-1 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| |||||
 Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 3012

RESULT 14

US-09-441-411-5
 ; Sequence 5, Application US/09441411
 ; Publication No. US20030008342A1
 ; GENERAL INFORMATION:

; APPLICANT: Scholler, Nathalie B.
 ; APPLICANT: Disis, Mary L.
 ; APPLICANT: Hellstrom, Ingegerd
 ; APPLICANT: Hellstrom, Karl Erik
 ; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
 ; FILE REFERENCE: 730033.409
 ; CURRENT APPLICATION NUMBER: US/09/441,411
 ; CURRENT FILING DATE: 1999-11-16
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 4473
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-441-411-5

Alignment Scores:
 Pred. No.: 0.00642 Length: 4473
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 11 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-441-411-5 (1-4473)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| |||||
 Db 3133 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 3186

RESULT 15

US-10-101-510-81
 ; Sequence 81, Application US/10101510
 ; Publication No. US20030148295A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WAN, JACKSON
 ; APPLICANT: WANG, YIXIN
 ; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
 ; FILE REFERENCE: 15117.0012
 ; CURRENT APPLICATION NUMBER: US/10/101,510
 ; CURRENT FILING DATE: 2002-03-20
 ; PRIOR APPLICATION NUMBER: 60/276,947
 ; PRIOR FILING DATE: 2001-03-20
 ; NUMBER OF SEQ ID NOS: 805
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 81
 ; LENGTH: 4473
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-101-510-81

Alignment Scores:
 Pred. No.: 0.00642 Length: 4473
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-10-101-510-81 (1-4473)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| |||||
 Db 3133 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 3186

Search completed: October 17, 2003, 03:53:11
 Job time : 35.0111 secs

THIS PAGE BLANK (USPTO)


```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 0.00745 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-08-625-101-1 (1-3768)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCACCCAGTCCCTTGGACAGCACCTTC 3012
RESULT 2
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3768
; OTHER INFORMATION: /note= "product = "cerb-b2"
```

```
US-08-356-786-1
Alignment Scores:
Pred. No.: 0.00745 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-08-356-786-1 (1-3768)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCACCCAGTCCCTTGGACAGCACCTTC 3012
RESULT 3
US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; TITLE OF INVENTION: HUMAN HER-2 EXPRESSION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748ris LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048,804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: NO
US-09-048-804-1

Alignment Scores:
Pred. No.: 0.0092 Length: 4473
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-048-804-1 (1-4473)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3133 GTGGTCATCCAGATGAGGACTTGGGCCACCCAGTCCCTTGGACAGCACCTTC 3186
RESULT 4
US-09-056-105-26
```

; Sequence 26, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
; APPLICANT: KIPPS, THOMAS J.
; APPLICANT: WU, YUNQI
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
; TITLE OF INVENTION: PROCESSING
; FILE REFERENCE: 233/221
; CURRENT APPLICATION NUMBER: US/09/056,105
; CURRENT FILING DATE: 1998-04-06
; EARLIER APPLICATION NUMBER: 60/043,467
; EARLIER FILING DATE: 1997-04-10
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-056-105-26

Alignment Scores:
Pred. No.: 0.0092 Length: 4473
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-056-105-26 (1-4473)

OY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| |||||
Db 3133 GTGGTCATCCAGAAATGAGGACTTGGCCCGCAGCCAGTCCCTTGGACAGCACCTTC 3186

RESULT 5

US-08-229-515A-9
; Sequence 9, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELEPHONE: 404-688-9880
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-08-229-515A (1-4530)
OY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| |||||
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGCCCGCAGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 7

US-09-167-322-4

US-08-229-515A-9

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-08-229-515A-9 (1-4530)

OY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| |||||
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGCCCGCAGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 6

US-08-645-865-9
; Sequence 9, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-645-865-9
Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-08-645-865-9 (1-4530)
OY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| |||||
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGCCCGCAGCCAGTCCCTTGGACAGCACCTTC 3162

; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavoragna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-167-322-4 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 8
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
US-09-527-487-1

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-527-487-1 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 9
US-09-877-177A-11
; Sequence 11, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-877-177A-11 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 10
US-08-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

; TOPOLOGY: linear
US-08-645-865-14

Alignment Scores:
Pred. No.: 0.0158 Length: 3955
Score: 81.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.62% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-08-645-865-14 (1-3955)
QY 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPhe 15
|||||
Db 1946 GAACGAGGCTGCCAGCAGCAGACAGACGACCGCGGTGACATTC 1990
|||||

RESULT 12
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3

Alignment Scores:
Pred. No.: 0.766 Length: 2385
Score: 68.00 Matches: 17
Percent Similarity: 66.67% Conservative: 3
Best Local Similarity: 56.67% Mismatches: 6
Query Match: 42.50% Indels: 4
DB: 2 Gaps: 2

JUNC_SEQ8_SEQ4 (1-30) x US-09-146-283-3 (1-2385)

```

```
Oy 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
; Sequence 3, Application US/08579823A
Db 1925 GACAAGGCGTGCCTCGCCGCGAGAGAGCCAGCCCTCTGACGTCCTCGAG----- 1975

Oy 21 GlyProAla---SerProLeuAspSerThr 29
; Sequence 3, Application US/08579823A
Db 1976 GCACCCGCGCTCGCCGCGAGAGAGCCAGCCAGCACA 2005

RESULT 13
US-08-579-823A-3
; Sequence 3, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-08-579-823A-3

Alignment Scores:
Pred. No.: 0.766 Length: 2385
Score: 68.00 Matches: 17
Percent Similarity: 66.67% Conservative: 3
Best Local Similarity: 56.67% Mismatches: 6
Query Match: 42.50% Indels: 4
DB: 3 Gaps: 2

JUNC_SEQ8_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)
Oy 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
; Sequence 3, Application US/08579823A
Db 1925 GACAAGGCGTGCCTCGCCGCGAGAGAGCCAGCCCTCTGACGTCCTCGAG----- 1975

Oy 21 GlyProAla---SerProLeuAspSerThr 29
; Sequence 3, Application US/08579823A
Db 1976 GCACCCGCGCTCGCCGCGAGAGAGCCAGCCAGCACA 2005

RESULT 15
US-08-776-251-3
; Sequence 3, Application US/08776251
```

```
RESULT 14
US-09-344-195-3
; Sequence 3, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-344-195-3

Alignment Scores:
Pred. No.: 0.766 Length: 2385
Score: 68.00 Matches: 17
Percent Similarity: 66.67% Conservative: 3
Best Local Similarity: 56.67% Mismatches: 6
Query Match: 42.50% Indels: 4
DB: 3 Gaps: 2

JUNC_SEQ8_SEQ4 (1-30) x US-09-344-195-3 (1-2385)
Oy 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
; Sequence 3, Application US/09344195
Db 1925 GACAAGGCGTGCCTCGCCGCGAGAGAGCCAGCCAGCACA 2005

Oy 21 GlyProAla---SerProLeuAspSerThr 29
; Sequence 3, Application US/09344195
Db 1976 GCACCCGCGCTCGCCGCGAGAGAGCCAGCCAGCACA 2005

RESULT 15
US-08-776-251-3
; Sequence 3, Application US/08776251
```

```

: Patent No. 6025340
: GENERAL INFORMATION:
: APPLICANT: Springer, Caroline J
: APPLICANT: Marais, Richard
: TITLE OF INVENTION: Surface expression of enzyme in gene directed prodruug therapy
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Nixon & Vanderhye
: STREET: 1100 No. 6025340th Glebe Road, 8th Floor
: CITY: Arlington
: STATE: Virginia
: COUNTRY: USA
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/776,251
: FILING DATE: 31-JAN-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/GB95/01782
: FILING DATE: 27-JUL-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9415167.7
: FILING DATE: 27-JUL-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Arthur R. Crawford
: REGISTRATION NUMBER: 25,327
: REFERENCE/DOCKET NUMBER: 620-20
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 153 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-776-251-3

```

```

Alignment Scores:
Pred. No.: 0.0516 Length: 153
Score: 66.00 Matches: 11
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 78.57% Mismatches: 0
Query Match: 41.25% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-08-776-251-3 (1-153)

Qy 1 GluArgGlyCysProAlaGluInArgAlaSerProValThr 14
      : : : : : : : : : : : : : : : : : : : :
Db 10 GACAAGGGGTGCCCGCCGAGCAGAGCAGCGCTCTGACG 51

```

Search completed: October 16, 2003, 17:16:36
Job time : 16.6049 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 28.4941 Seconds
(without alignments)
2842.104 Million cell updates/sec

Title: JUNC_SEQ8_SEQ4

Perfect score: 160
Sequence: 1 ERGCPAEQASPVTFQNEGLGPASPLDSTF 30

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/cn2.1/USPTO_SPOOL_P/HOLLERAN480/runat_15102003_131912_20526/app_query.fasta_1.4685
-DB=N.Geneseq_19jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptm -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGN 1.1.1758 -runat_15102003_131912_20526 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	147	91.9	2763	24	ABA92252	Mouse Her-2/neu ex
2	147	91.9	2781	24	ABA92253	Her-2/neu extracel
3	83	51.9	1115	21	AAF21778	Human breast and o
4	83	51.9	1755	24	AAD32746	Human cDNA for the
5	83	51.9	1767	24	AAD32744	Human cDNA for the
6	83	51.9	1773	24	AAD32747	Human cDNA for the
7	83	51.9	1806	24	AAD32745	Human cDNA for the
8	83	51.9	3600	21	AAA89736	Human HER-2/neu co
9	83	51.9	3678	24	ABK86207	cDNA encoding huma
10	83	51.9	3768	17	AAT40739	HER-2/neu oncogene
11	83	51.9	3768	20	AAK01912	Human HER-2/neu on
12	83	51.9	3768	21	AAA09455	Human heregulin 2
13	83	51.9	3768	22	AAH23392	Human HER-2/neu pr
14	83	51.9	3768	24	ABZ35744	Human ERBB2 polynu
15	83	51.9	3768	24	ABX0987	Human ERBB2 DNA fr
16	83	51.9	3768	24	AAD43935	Human HER-2 CDNA.
17	83	51.9	3768	24	AAD43986	Human Her2 antigen
18	83	51.9	3768	24	ABV78168	Human ERBB2 DNA SE
19	83	51.9	3768	24	AAD32743	Human Her-2/neu pr
20	83	51.9	3768	24	ABA92250	Human Her-2/neu CD
21	83	51.9	3768	24	ABK10730	Human Her-2/neu DN
22	83	51.9	3768	24	ABL91709	Human polynucleoti
23	83	51.9	3768	24	ABK14058	Human HER2 (Erbb2)
24	83	51.9	4299	14	AAQ46083	Sequence encoding
25	83	51.9	4472	21	AAA14812	cDNA encoding the
26	83	51.9	4473	19	ABQ76220	Human tumour antiq
27	83	51.9	4473	20	AZ331071	Human Her-2 nucleic acid
28	83	51.9	4473	24	ABZ39969	Human gene express
29	83	51.9	4473	24	AAD38904	Human Her-2 DNA.
30	83	51.9	4530	16	AAT01585	Her-2/neu (ERBB2/c
31	83	51.9	4530	18	AAT71253	Human Her-2 gene.
32	83	51.9	4530	21	AZ60815	Nucleotide sequenc
33	83	51.9	4530	22	AD19731	Human tyrosine kin
34	83	51.9	4530	24	ABZ35012	Human gene express
35	83	51.9	4530	24	ABV94128	Breast carcinoma r
36	83	51.9	4530	24	ABN85585	Human HER2-neu SQ
37	83	51.9	4530	24	ABK83918	Human cDNA differe
38	83	51.9	4530	25	ACC50139	Breast cancer asso
39	83	51.9	4530	25	ABQ83856	Human Her2/Neu enc
40	83	51.9	9274	22	AAF24297	HER2 transgene pla
41	83	51.9	9274	24	AAD43934	HER-2 transgene pl
42	83	51.9	9274	24	ABK14057	Human HER2 (Erbb2)
43	81	50.6	3771	21	AAA89737	Mouse Her-2/neu CD
44	81	50.6	3771	22	AAA42210	Nucleotide sequenc
45	81	50.6	3771	24	ABA92251	Mouse Her-2/neu CD

ALIGNMENTS

RESULT 1
ABA92252
ID ABA92252 standard; cDNA; 2763 BP.
XX
AC ABA92252;
XX
DT 17-JUN-2002 (first entry)
XX
DE Mouse Her-2/neu extracellular-phosphorylation domain fusion cDNA.
XX
KW Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
XX
KW receptor; mouse; gene therapy; gene; ss.
XX
OS Mus musculus.
XX
FH Key
FT CDS Location/Qualifiers
1..2763

QY 1 GluArgGlyCysProAlaGluInArgAlaSerProValThrPheGlnAsnGluAspLeu 20
 Db 1918 GAACGAGGTGCCAGCAGCAGAGAGCCAGTGCCTCTCAGAACGAGGACTTA 1977

QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
 Db 1978 GGCCCTCCAGCCCATGGACACGACCTTC 2007

RESULT 3
 AAF21778
 ID AAF21778 standard; DNA; 1115 BP.
 XX AAF21778;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Human breast and ovarian cancer associated antigen gene SEQ ID 165.
 XX
 KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antinflammatory; antiulcer; vulnery; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200055173-A1.
 PN 21-SEP-2000.
 PD
 XX
 XX 08-MAR-2000; 2000WO-US05881.
 PF
 XX
 XX 12-MAR-1999; 99US-0124270.
 PR
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2000-611515/58.
 DR P-PSDB; AAB58875.
 DR
 XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases -
 XX
 XX Claim 1; Page 604; 1299pp; English.
 PS
 XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antidiabetic; antinflammatory; antiulcer; vulnery; anticonvulsant;
 CC antibacterial; antifungal; antiparasitic and cardiant activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and agonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemias; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 XX
 XX Sequence 1115 BP; 210 A; 336 C; 338 G; 222 T; 9 other;

Alignment Scores:
 Pred. No.: 0.0139 Length: 1115
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 21 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AAF21778 (1-1115)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 Db 446 GTGGTCATCCAGAAATGAGGACTTGGCCAGCCAGTCCCTTGGACAGCACCTTC 499

RESULT 4

AAD32746
 ID AAD32746 standard; cDNA; 1755 BP.
 XX

AC AAD32746;
 XX

DT 01-JUL-2002 (first entry)
 XX

DE Human cDNA for the clone HICD_native_coding_region.
 XX

KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
 XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT CDS 1..1755

FT /*tag= a
 FT /product= "Human protein encoded by cDNA for the clone
 FT HICD_native_coding_region"
 FT /transl_except= (pos:1741..1752, aa:Leu-Glu)
 FT /note= "CDS does not include stop codon"
 FT /partial

XX WO200214503-A2.
 XX

PD 21-FEB-2002.
 XX

PF 14-AUG-2001; 2001WO-US41733.
 XX

PR 14-AUG-2000; 2000US-225152P.
 XX

PR 28-SEP-2000; 2000US-236428P.
 XX

PR 21-FEB-2001; 2001US-270520P.
 XX

PA (CORI-) CORIXA CORP.
 XX

PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
 PI McNeill PD, Vedvick TS;
 XX

DR WPI; 2002-280758/32.
 XX

XX P-PSDB; AAE20483.
 XX
 PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,
 PT prevention and diagnosis of cancer, preferably breast cancer -

XX Example 5; Page 118-119; 129pp; English.

XX The invention relates to an isolated Her-2/Neu polypeptide composition
 CC effective for eliciting an immune response. The invention is useful for
 CC eliciting an immune response in a patient, where the patient is human
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
 CC The composition is useful for the therapy and diagnosis of cancer,
 CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
 CC and other compositions for the diagnosis, prevention and treatment of
 CC human malignancies, for stimulating and/or expanding T cells specific for
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
 CC patient. The invention is useful for stimulating a T cell response in a
 CC human patient, as probe or primer for nucleic acid hybridisation, to

CC selectively form duplex molecules with complementary stretches of the
 CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
 CC length gene from a suitable library, and to direct expression of a
 CC polypeptide in appropriate host cells. The composition is useful in
 CC prophylactic or therapeutic applications and for the treatment of cancer,
 CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
 CC associated malignancies. The invention is useful in gene therapy. The
 CC present sequence is human cDNA for the clone HICD_native_coding_region.
 XX
 SQ Sequence 1755 BP; 376 A; 517 C; 531 G; 331 T; 0 other;

Alignment Scores:
 Pred. No.: 0.0239 Length: 1755
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 24 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AAD32746 (1-1755)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 937 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCAGCTTC 990

RESULT 5
 AAD32744
 ID AAD32744 standard; cDNA; 1767 BP.
 XX
 AC AAD32744;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Human cDNA for the clone HICD_CT_His_coding_region.
 XX
 KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1764
 FT /*tag= a
 FT /product= "Human protein encoded by cDNA for the clone
 HICD_CT_His_coding_region"
 XX
 PN W0200214503-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 14-AUG-2001; 2001WO-US41733.
 XX
 PR 14-AUG-2000; 2000US-225152P.
 PR 28-SEP-2000; 2000US-236428P.
 PR 21-FEB-2001; 2001US-270520P.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
 PI McNeill PD, Vedvick TS;
 XX
 DR WPI; 2002-280758/32.
 DR P-PSDB; AAE20481.
 XX
 PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,
 PT prevention and diagnosis of cancer, preferably breast cancer -
 XX
 PS Example 5; Page 117-118; 129pp; English.
 XX
 CC The invention relates to an isolated Her-2/Neu polypeptide composition
 CC effective for eliciting an immune response. The invention is useful for
 CC eliciting an immune response in a patient, where the patient is human
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.

CC The composition is useful for the therapy and diagnosis of cancer,
 CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
 CC and other compositions for the diagnosis, prevention and treatment of
 CC human malignancies, for stimulating and/or expanding T cells specific for
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
 CC patient. The invention is useful for stimulating a T cell response in a
 CC human patient, as probe or primer for nucleic acid hybridisation, to
 CC selectively form duplex molecules with complementary stretches of the
 CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
 CC length gene from a suitable library, and to direct expression of a
 CC polypeptide in appropriate host cells. The composition is useful in
 CC prophylactic or therapeutic applications and for the treatment of cancer,
 CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
 CC associated malignancies. The invention is useful in gene therapy. The
 CC present sequence is human cDNA for the clone HICD_CT_His_coding_region.
 XX
 SQ Sequence 1767 BP; 381 A; 521 C; 529 G; 336 T; 0 other;

Alignment Scores:
 Pred. No.: 0.0241 Length: 1767
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 24 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AAD32744 (1-1767)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 937 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCAGCTTC 990

RESULT 6
 AAD32747
 ID AAD32747 standard; cDNA; 1773 BP.
 XX
 AC AAD32747;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Human cDNA for the clone HICD_in_ppDM_coding_sequence.
 XX
 KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1770
 FT /*tag= a
 FT /product= "Human protein encoded by cDNA for the clone
 HICD_in_ppDM_coding_sequence"
 XX
 PN W0200214503-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 14-AUG-2001; 2001WO-US41733.
 XX
 PR 14-AUG-2000; 2000US-225152P.
 PR 28-SEP-2000; 2000US-236428P.
 PR 21-FEB-2001; 2001US-270520P.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
 PI McNeill PD, Vedvick TS;
 XX
 DR WPI; 2002-280758/32.
 DR P-PSDB; AAE20481.
 XX
 PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,
 PT prevention and diagnosis of cancer, preferably breast cancer -

XX Example 5; Page 119; 129pp; English.

PS The invention relates to an isolated Her-2/Neu polypeptide composition

XX effective for eliciting an immune response. The invention is useful for

CC eliciting an immune response in a patient, where the patient is human

CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.

CC The composition is useful for the therapy and diagnosis of cancer,

CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine

CC and other compositions, for the diagnosis, prevention and treatment of

CC human malignancies, for stimulating and/or expanding T cells specific for

CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a

CC patient. The invention is useful for stimulating a T cell response in a

CC human patient, as probe or primer for nucleic acid hybridisation, to

CC selectively form duplex molecules with complementary stretches of the

CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full

CC length gene from a suitable library, and to direct expression of a

CC polypeptide in appropriate host cells. The composition is useful in

CC prophylactic or therapeutic applications and for the treatment of cancer,

CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-

CC associated malignancies. The invention is useful in gene therapy. The

CC present sequence is human cDNA for the clone HICD_in_pPDM_coding_

XX sequence.

SQ Sequence 1773 BP; 383 A; 528 C; 530 G; 332 T; 0 other;

Alignment Scores:

Pred. No.:	0.0242	Length:	1773
Score:	83.00	Matches:	16
Percent Similarity:	88.8%	Conservative:	0
Best Local Similarity:	88.8%	Mismatches:	2
Query Match:	51.88%	Indels:	0
DB:	24	Gaps:	0

JUNC_SEQ8_SEQ4 (1-30) x AAD32747 (1-1773)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 961 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 1014

RESULT 7

AAD32745

ID AAD32745 standard; cDNA; 1806 BP.

XX AAD32745;

AC AAD32745;

DT 01-JUL-2002 (first entry)

XX Human cDNA for the clone HICD_plus_8_HIS.

DE Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;

XX human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1803

FT /*tag= a

FT /product= "Human protein encoded by cDNA for the clone

FT HICD_plus_8_HIS"

FT /transl_except= (pos:1543..1545, aa:Pro)

XX WO200214503-A2.

XX 21-FEB-2002.

XX 14-AUG-2001; 2001WO-US41733.

XX 14-AUG-2000; 2000US-225152P.

PR 28-SEP-2000; 2000US-236428P.

PR 21-FEB-2001; 2001US-270520P.

XX (CORI-) CORIXA CORP.

XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;

PI McNeill PD, Vedrick TS;

XX WPI; 2002-280758/32.

DR P-PSDB; AAE20482.

XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,

PT prevention and diagnosis of cancer, preferably breast cancer -

XX Example 5; Page 118; 129pp; English.

XX The invention relates to an isolated Her-2/Neu polypeptide composition

CC effective for eliciting an immune response. The invention is useful for

CC eliciting an immune response in a patient, where the patient is human

CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.

CC The composition is useful for the therapy and diagnosis of cancer,

CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine

CC and other compositions, for the diagnosis, prevention and treatment of

CC human malignancies, for stimulating and/or expanding T cells specific for

CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a

CC patient. The invention is useful for stimulating a T cell response in a

CC human patient, as probe or primer for nucleic acid hybridisation, to

CC selectively form duplex molecules with complementary stretches of the

CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full

CC length gene from a suitable library, and to direct expression of a

CC polypeptide in appropriate host cells. The composition is useful in

CC prophylactic or therapeutic applications and for the treatment of cancer,

CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-

CC associated malignancies. The invention is useful in gene therapy. The

CC present sequence is human cDNA for the clone HICD_plus_8_HIS.

XX Sequence 1806 BP; 391 A; 530 C; 544 G; 341 T; 0 other;

Alignment Scores:

Pred. No.:	0.0248	Length:	1806
Score:	83.00	Matches:	16
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	2
Query Match:	51.88%	Indels:	0
DB:	24	Gaps:	0

JUNC_SEQ8_SEQ4 (1-30) x AAD32745 (1-1806)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 994 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 1047

RESULT 8

AAA89736

ID AAA89736 standard; DNA; 3600 BP.

XX AAA89736;

AC AAA89736;

DT 12-JAN-2001 (first entry)

XX Human HER-2/neu coding sequence.

DE Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;

XX breast cancer; prostate cancer; ovarian cancer; lung cancer;

XX colon cancer; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..3600

FT /*tag= a

FT /product= "HER-2/neu protein"

XX WO2000044899-A1.

XX 03-AUG-2000.

PF 28-JAN-2000; 2000WO-US02164.
 XX
 PR 29-JAN-1999; 99US-0117976.
 XX
 PA (CORI-) CORIXA CORP.
 PA (SMK) SMITHKLINE BEECHAM.
 XX
 PI Cheever MA, Gheysen D;
 XX
 DR WPI; 2000-505976/45.
 DR P-PSDB; RAB21198, RAB21208.
 XX
 PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
 PT useful for vaccinating against breast, ovarian, colon, lung and
 PT prostate cancers -
 XX
 PS Disclosure; Fig 15; 128pp; English.
 XX
 CC The present sequence encodes the human HER-2/neu protein. HER-2/neu is
 CC a member of the tyrosine kinase family of receptor-like glycoproteins
 CC and shows homology to the epidermal growth factor receptor (EGFR). It
 CC probably plays a part in cell growth and/or differentiation. The
 CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising
 CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
 CC domain may be used to treat or prevent cancer by eliciting or
 CC enhancing an immune response to the HER-2/neu protein. It may be used
 CC to treat malignancies such as breast, ovarian, colon, lung and
 CC prostate cancers, and may be used as an antigen to vaccinate against
 CC these neoplasias.
 XX
 SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.0566 Length: 3600
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 21 Gaps: 0
 JUNC_SEQ8_SEQ4 (1-30) x AAB89736 (1-3600)
 QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 Db 2959 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3012
 RESULT 9
 ABK86207
 ID ABR86207 standard; cDNA; 3678 BP.
 AC ABK86207;
 XX
 DT 24-SEP-2002 (first entry)
 XX
 DE cDNA encoding human breast cancer antigen, Her2 variant.
 XX
 KW Human; Her2; cytostatic; antiviral; immunostimulant;
 KW cell-mediated immune response; tumour; breast cancer;
 KW virus infection; prostate cancer; colorectal cancer; pancreatic cancer;
 KW lymphoma; leukaemia; hepatitisvirus; lentivirus; herpesvirus;
 KW human immunodeficiency virus; HIV; flavivirus; pestivirus; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 7..3678
 FT /*tag= a
 FT /product= "Breast cancer antigen Her2 variant"
 XX
 PN W0200240059-A2.
 XX
 PD 23-MAY-2002.
 XX

PF 01-NOV-2001; 2001WO-US45626.
 XX
 PR 01-NOV-2000; 2000US-0704232.
 XX
 PA (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.
 PA (MINC/) MINCHEFF M S.
 PA (LOUK/) LOUKINOV D I.
 PA (ZOUB/) ZOUBAK S.
 XX
 PI Mincheff MS, Loukinov DI, Zoubak S;
 XX
 DR WPI; 2002-527524/56.
 DR P-PSDB; AAU98923.
 XX
 PT Inducing a cell-mediated immune response against a target antigen,
 PT reducing undesired cells and stimulating presentation of an antigen by
 PT a cell, comprises administering a polynucleotide encoding a variant of
 PT an antigen -
 XX
 PS Disclosure; Page 128-134; 146pp; English.
 XX
 CC The invention relates to a method of inducing a cell-mediated immune
 CC response against a cell comprising a target antigen (I) in a subject,
 CC treating a subject having undesired cells, for example tumour cells
 CC or virally infected cells (C), reducing the number of (C) in a subject,
 CC and stimulating presentation of (I) by a cell. This is done by
 CC administering a polynucleotide (II) encoding a variant of (I), so that
 CC (II) expressed in a cell and cell-mediated immune response is induced.
 CC The method can be used to treat prostate cancer, breast cancer,
 CC colorectal cancer and pancreatic cancer, as well as lymphomas and
 CC leukaemias. The method is also useful in treating chronic viral
 CC infections such as those caused by hepatitisviruses, lentiviruses
 CC (including human immunodeficiency virus (HIV)), herpesviruses and the
 CC flaviviruses and pestiviruses. The present sequence represents the coding
 CC sequence of human breast cancer antigen, Her2 variant, used as a target
 CC antigen in the method of the invention.
 XX
 SQ Sequence 3678 BP; 746 A; 1137 C; 1091 G; 704 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.0581 Length: 3678
 Score: 83.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 51.88% Indels: 0
 DB: 24 Gaps: 0
 JUNC_SEQ8_SEQ4 (1-30) x ABR86207 (1-3678)
 QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
 Db 2869 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 2922
 RESULT 10
 AAT40739
 ID AAT40739 standard; cDNA; 3768 BP.
 XX
 AC AAT40739;
 XX
 DT 01-JAN-1997 (first entry)
 XX
 DE HER-2/neu oncogene.
 XX
 KW HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
 KW breast cancer; ovary cancer; colon cancer; lung cancer;
 KW prostate cancer; genetic immunisation; tumour; vaccine; vector;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3765
 FT /*tag= b
 FT

```

FT      /note= "nucleotides 2026-3765 (claim 1) code for
FT      HER-2/neu intracellular domain"
XX      WO9630514-A1.
PN      03-OCT-1996.
XX      28-MAR-1996; 96WO-US01689.
XX      31-MAR-1995; 95US-0414417.
XX      (UNIW ) UNIV WASHINGTON.
XX      Cheever MA, Disis ML;
XX      WPI; 1996-455361/45.
XX      P-PSDB; AAW01111.
XX      DNA encoding HER-2-neu poly:peptide(s) - used for prevention or
XX      treatment of malignancies with which the HER-2/neu oncogene is
XX      associated
XX      Claim 1; Page 49-56; 71pp; English.
XX      Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or
XX      c-erbB2) protein (AAW01111). The oncogene is overexpressed in various
XX      cancers, including breast, ovarian, colon, lung and prostate, and
XX      appears to induce malignancies through quantitative mechanisms that
XX      result from increased or deregulated expression of an essentially
XX      normal gene product. Nucleotides 2026-3765 of the cDNA sequence
XX      code for the intracellular domain (Lys676-Val1255) of the HER-2/neu
XX      protein, which is useful for immunisation against malignancy.
XX      Nucleic acids can be used to direct expression of the intracellular
XX      domain in transformed host cells, or are used, alone or in a viral
XX      vector, for genetic immunisation of an animal.
XX      Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
SQ      Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 17 Gaps: 0
JUNC_SEQ8_SEQ4 (1-30) x AAT40739 (1-3768)
QY      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db      2959 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3012
RESULT 11
AAAX01912
ID      AAX01912 standard; DNA; 3768 BP.
XX      AAX01912;
XX      21-APR-1999 (first entry)
XX      Human HER-2/neu oncogene DNA.
XX      HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
XX      malignancy; treatment; tumour; ss.
XX      Homo sapiens.
XX      Key Location/Qualifiers
XX      CDS 1..3768
XX      FT /*tag= a
XX      FT /product= "HER-2/neu"
XX      FT /note= "oncogene"
XX      FT misc_feature 2026..3765

```

```

FT      /*tag= b
FT      /note= "region which elicits immune response"
XX      US5869445-A.
PN      09-FEB-1999.
XX      01-APR-1996; 96US-0625101.
XX      01-APR-1996; 96US-0625101.
XX      17-MAR-1993; 93US-0033644.
XX      12-AUG-1993; 93US-0106112.
XX      31-MAR-1995; 95US-0414417.
XX      (UNIW ) UNIV WASHINGTON.
XX      Cheever MA, Disis ML;
XX      WPI; 1999-152835/13.
XX      P-PSDB; AAW92406.
XX      Use of HER-2/neu polypeptides - for eliciting an immune response to
XX      an HER-2/neu associated malignancy, particularly for treating or
XX      preventing tumours
XX      Claim 1a; Column 23-32; 26pp; English.
XX      This sequence encodes the human HER-2/neu oncogene protein. A fragment
XX      of this protein is used in a method for eliciting or enhancing an immune
XX      response to HER-2/neu protein. The polypeptide can stimulate T cells and
XX      B cells to produce an immune response to the HER-2/neu protein. The
XX      method can be used for immunisation against a malignancy in which the
XX      HER-2/neu oncogene is associated and in the treatment of an existing
XX      tumour, or to prevent tumour occurrence or reoccurrence.
XX      Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
SQ      Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 20 Gaps: 0
JUNC_SEQ8_SEQ4 (1-30) x AAX01912 (1-3768)
QY      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db      2959 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3012
RESULT 12
AAAX09455
ID      AAX09455 standard; DNA; 3768 BP.
XX      AAX09455;
XX      10-AUG-2000 (first entry)
XX      Human heregulin 2 (Her2) coding sequence.
XX      Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
XX      self-protein; cell-associated peptide antigen; foreign epitope;
XX      cancer; breast cancer; prostate cancer; ss.
XX      Homo sapiens.
XX      OS
XX      WO200020027-A2.
XX      13-APR-2000.
XX      PD
XX      05-OCT-1999; 99WO-DK00525.
XX      FT

```

```

PR 05-OCT-1998; 98DK-0001261.
PI 20-OCT-1998; 98US-0105011.
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
DR P-PSDB; AAY92620.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
XX Claim 62; Page 187-193; 220pp; English.
XX
XX The claims detail a method for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, human
CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively.
XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AAA09455 (1-3768)
Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 13
AAH23392
ID AAH23392 standard; DNA; 3768 BP.
XX
XX AAH23392;
XX
XX 25-SEP-2001 (first entry)
XX
XX Human HER-2/neu protein encoding DNA.
XX
XX Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1..3768
CDS /*tag= a
FT /product= "HER-2/neu protein"
XX
XX WO200153463-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US01850.

```

```

XX 21-JAN-2000; 2000US-0177545.
XX (CORI-) CORIXA CORP.
XX
XX Cheever MA, Hand-Zimmermann S;
PI
XX WPI; 2001-476112/51.
DR P-PSDB; AAB85458.
XX
XX New antigen-presenting cells, useful as vaccines for eliciting or
PT enhancing an immune response to HER-2/neu protein, particularly useful
PT for treating or preventing cancer, e.g. breast cancer
XX
XX Claim 1; Page 41-46; 49pp; English.
XX
XX The invention provides an isolated antigen-presenting cell, which
CC expresses at least an immunogenic portion of a polypeptide that produces
CC an immune response to HER-2/neu protein. The antigen-presenting cells are
CC useful as vaccines for eliciting or enhancing an immune response to
CC HER-2/neu protein, particularly in treating or preventing malignancies in
CC which the HER-2/neu oncogene is associated. Specifically, these are
CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,
CC colon, lung or prostate cancers. The present sequence represents a DNA
CC encoding the human HER-2/neu protein (also known as p185 or c-erbB2).
XX
SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 22 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AAH23392 (1-3768)
Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 14
ABZ35744
ID ABZ35744 standard; DNA; 3768 BP.
XX
XX AC ABZ35744;
XX
XX 07-FEB-2003 (first entry)
XX
XX Human ERBB2 polynucleotide SEQ ID NO 52.
XX
XX Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
KW protozoacide; gene expression; antisense; tumour; infection; Plasmodium;
KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
KW Hepatitis C virus; human papilloma virus; gene; ds.
XX
XX Homo sapiens.
XX
XX PN DE10100588-A1.
XX
XX 18-JUL-2002.
XX
XX 09-JAN-2001; 2001DE-1000588.
XX
XX 09-JAN-2001; 2001DE-1000588.
XX
XX (RIBO-) RIBOPHARMA AG.
XX
XX Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
XX WPI; 2002-683450/74.
XX

```

PT Inhibiting expression of target genes, useful e.g. for treating tumors,
PT by introducing into cells two double-stranded RNAs that are
PT complementary to the target -
XX
XX Claim 13; Page 38-39; 100pp; German.
XX
CC The invention relates to inhibiting expression of a target gene in a cell
CC by introducing at least two oligoribonucleotides (dsRNAI and II), both
CC with a double-stranded (ds) structure of at most 49 sequential nucleotide
CC pairs. At least part of one strand (S1, S2) of the ds structures in each
CC of dsRNAI and II are complementary to regions in the target gene. The
CC method uses antisense inhibition of gene expression using double stranded
CC RNA inhibition (RNAi). The method is particularly used to treat tumours
CC or infections, especially by Plasmodium or viruses/viroids (pathogenic on
CC humans, animals or plants). The method provides more effective inhibition
CC of expression than known methods using a single dsRNA, even at very low
CC concentrations. When dsRNA has at least one unpaired nucleotide at the
CC end, stability (and thus effective concentration in the cell) is
CC improved and efficiency can be increased further by pretreating the cells
CC with interferon. The present sequence is that of a target DNA of the
CC invention.
XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 24 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x ABZ35744 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 15
ABX09987
ID ABX09987 standard; DNA; 3768 BP.
XX
AC ABX09987;
XX
DT 23-JAN-2003 (first entry)
XX
DE Human ERBB2 DNA fragment SEQ ID 52.
XX
KW Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;
KW prion; inhibition; human; ds.
XX
OS Homo sapiens.
XX
PN DE10100587-C1.
XX
PD 21-NOV-2002.
XX
PF 09-JAN-2001; 2001DE-1000587.
XX
PR 09-JAN-2001; 2001DE-1000587.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-742209/81.
XX
PT Inhibiting expression of target genes, e.g. oncogenes, in cells, by
PT introduction of complementary double-stranded oligoribonucleotide,
PT after treating the cell with interferon -
XX
PS Disclosure; Page 43-44; 98pp; German.
XX

CC This invention describes a novel method for inhibiting expression of a
CC target gene by introducing into the cell that contains the target gene
CC at least one oligoribonucleotide (dsRNAI) that has a double-stranded
CC (ds) structure of not more than 49 consecutive nucleotides (nt), where
CC at least a segment of one strand of the ds structure is complementary
CC with the target gene and the cells are treated with interferon before
CC introduction of dsRNAI. The method is used to inhibit expression of
CC target genes, particularly oncogenes, cytokine genes, Id (not defined)
CC protein genes; developmental or prion genes, or genes expressed in
CC pathogenic organisms (particularly plasmodia) or in viruses or viroids
CC (pathogenic in humans, animals or plants). Treating the cells with
CC interferon greatly increases the extent to which dsRNA can inhibit
CC expression of the target genes, and the effect is even greater when dsRNA
CC are modified to increase their stability. ABX09936-ABX10075 represent
CC gene fragments used to illustrate the method of the invention.
XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 24 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x ABX09987 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012

Search completed: October 15, 2003, 23:48:48
Job time : 31.4941 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:09:13 : Search time 432.026 Seconds
(without alignments)
2840.777 Million cell updates/sec

Title: JUNC_SEQ8_SEQ4
Perfect score: 160
Sequence: 1 ERGCPAEQRASPVTFQNDLGPASPLDSTF 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+p2n.model -DEV=rlp
-Q=/cgn2_1/USPTO.spool_p/HOLLERAM480/runat_15102003_131912_20536/app_query.fasta_1.4685
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAM480 -CGN_1_1_22724 -runat_15102003_131912_20536 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	gb.ba.*	GenEmbl.*
2:	gb.htg.*	
3:	gb.in.*	
4:	gb.om.*	
5:	gb.ov.*	
6:	gb.pat.*	
7:	gb.ph.*	
8:	gb.pl.*	
9:	gb.pr.*	
10:	gb.ro.*	
11:	gb.sts.*	
12:	gb.sy.*	
13:	gb.un.*	
14:	gb.vi.*	
15:	em.ba.*	
16:	em.fun.*	
17:	em.hum.*	
18:	em.in.*	
19:	em.mu.*	
20:	em.om.*	
21:	em.or.*	
22:	em.ov.*	
23:	em.pat.*	
24:	em.ph.*	
25:	em.pl.*	
26:	em.ro.*	
27:	em.sts.*	
28:	em.un.*	

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	147	91.9	2763	6	AX380942	AX380942 Sequence
2	147	91.9	2781	6	AX380944	AX380944 Sequence
3	83	51.9	1755	6	AX384609	AX384609 Sequence
4	83	51.9	1767	6	AX384607	AX384607 Sequence
5	83	51.9	1773	6	AX384610	AX384610 Sequence
6	83	51.9	1806	6	AX384608	AX384608 Sequence
7	83	51.9	3678	6	AX505114	AX505114 Sequence
8	83	51.9	3768	6	AR034479	AR034479 Sequence
9	83	51.9	3768	6	AX060704	AX060704 Sequence
10	83	51.9	3768	6	AX201817	AX201817 Sequence
11	83	51.9	3768	6	AX380923	AX380923 Sequence
12	83	51.9	3768	6	AX384604	AX384604 Sequence
13	83	51.9	3768	6	AX465456	AX465456 Sequence
14	83	51.9	3768	6	AX467229	AX467229 Sequence
15	83	51.9	3768	6	AX481438	AX481438 Sequence
16	83	51.9	3780	4	AB008451	AB008451 Canis fam
17	83	51.9	4473	6	AR080259	AR080259 Sequence
18	83	51.9	4473	6	AR167390	AR167390 Sequence
19	83	51.9	4473	9	HSEB2R	X03363 Human c-erb
20	83	51.9	4530	6	AR202597	AR202597 Sequence
21	83	51.9	4530	6	AR283481	AR283481 Sequence
22	83	51.9	4530	6	AX282577	AX282577 Sequence
23	83	51.9	4530	6	AX587649	AX587649 Sequence
24	83	51.9	4530	6	AX644071	AX644071 Sequence
25	83	51.9	4530	6	BD005474	BD005474 Cellular
26	83	51.9	4530	6	I21124	I21124 Sequence 9
27	83	51.9	4530	6	I59745	I59745 Sequence 9
28	83	51.9	4530	9	HUMHER2A	M11730 Human tyros
29	83	51.9	9274	6	AX060703	AX060703 Sequence
30	81	50.6	3771	6	AX189662	AX189662 Sequence
31	81	50.6	3771	6	AX380925	AX380925 Sequence
32	81	50.6	3955	6	AX380924	AX380924 Sequence
33	81	50.6	3955	6	I21129	I21129 Sequence 14
34	81	50.6	3955	6	I59750	I59750 Sequence 14
35	81	50.6	3955	10	RNNEUR	X03362 Rat mRNA fo
36	81	50.6	4694	10	BC046811	BC046811 Mus muscu
37	81	50.6	4695	10	BC053078	BC053078 Mus muscu
38	81	50.6	4727	10	AY116182	AY116182 Rattus no
39	80	50.0	4062	10	HAMNEU	D16295 Mesocricetu
40	79	49.4	13450	9	AB096612	AB096612 Homo sapi
41	79	49.4	16572	9	AB096614	AB096614 Homo sapi
42	79	49.4	20271	9	AB096613	AB096613 Homo sapi
43	79	49.4	30837	9	AY208911	AY208911 Homo sapi
44	79	49.4	161815	9	AC079199	AC079199 Homo sapi
45	79	49.4	168585	9	AC040933	AC040933 Homo sapi

ALIGNMENTS

RESULT 1


```
AX380942
LOCUS AX380942 2763 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 28 from Patent WO0212341.
ACCESSION AX380942
VERSION AX380942.1 GI:19575786
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
AUTHORS Cheever, M.A. and Gheysen, D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 28 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
1. 2763
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD fusion protein cDNA"
BASE COUNT 571 a 855 c 772 g 565 t
ORIGIN
Alignment Scores:
Pred. No.: 3.46e-11 Length: 2763
Score: 147.00 Matches: 27
Percent Similarity: 96.67% Conservative: 2
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 91.88% Indels: 0
Gaps: 0
JUNC_SEQ8_SEQ4 (1-30) x AX380942 (1-2763)
Qy 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACAGAGCCAGCCAGTGCAGTCTCAGACGAGGACTTA 1977
Qy 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCCCTCCAGCCCATGGACGACCTTC 2007
RESULT 2
AX380944
LOCUS AX380944 2781 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 30 from Patent WO0212341.
ACCESSION AX380944
VERSION AX380944.1 GI:19575787
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
AUTHORS Cheever, M.A. and Gheysen, D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 30 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
1. 2781
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD-TcP0 fusion protein cDNA"
BASE COUNT 574 a 859 c 779 g 569 t
ORIGIN
Alignment Scores:
Pred. No.: 3.49e-11 Length: 2781
Score: 147.00 Matches: 27
Percent Similarity: 96.67% Conservative: 2
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 91.88% Indels: 0
Gaps: 0
JUNC_SEQ8_SEQ4 (1-30) x AX380942 (1-2763)
Qy 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACAGAGCCAGCCAGTGCAGTCTCAGACGAGGACTTA 1977
Qy 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCCCTCCAGCCCATGGACGACCTTC 2007
RESULT 2
AX380944
LOCUS AX380944 2781 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 30 from Patent WO0212341.
ACCESSION AX380944
VERSION AX380944.1 GI:19575787
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
AUTHORS Cheever, M.A. and Gheysen, D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 30 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
1. 2781
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD-TcP0 fusion protein cDNA"
BASE COUNT 574 a 859 c 779 g 569 t
ORIGIN
Alignment Scores:
Pred. No.: 3.49e-11 Length: 2781
Score: 147.00 Matches: 27
Percent Similarity: 96.67% Conservative: 2
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 91.88% Indels: 0
Gaps: 0
JUNC_SEQ8_SEQ4 (1-30) x AX380942 (1-2763)
Qy 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACAGAGCCAGCCAGTGCAGTCTCAGACGAGGACTTA 1977
Qy 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCCCTCCAGCCCATGGACGACCTTC 2007
```

```
JUNC_SEQ8_SEQ4 (1-30) x AX380944 (1-2781)
Qy 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACAGAGCCAGCCAGTGCAGTCTCAGACGAGGACTTA 1977
Qy 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCCCTCCAGCCCATGGACGACCTTC 2007
RESULT 3
AX384609
LOCUS AX384609 1755 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 6 from Patent WO0214503.
ACCESSION AX384609
VERSION AX384609.1 GI:19577810
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
TITLE Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 6 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
1. 1755
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 376 a 517 c 531 g 331 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0303 Length: 1755
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
Gaps: 0
JUNC_SEQ8_SEQ4 (1-30) x AX384609 (1-1755)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 937 GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGTCCCTTGGACAGACCTTC 990
RESULT 4
AX384607
LOCUS AX384607 1767 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 4 from Patent WO0214503.
ACCESSION AX384607
VERSION AX384607.1 GI:19577808
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
TITLE Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 4 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
1. 1767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

BASE COUNT 381 a 521 c 529 g 336 t
ORIGIN

Alignment Scores:
Pred. No.: 0.0305 Length: 1767
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AX384607 (1-1767)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 937 GTGGTCATCCAGAAAGAGGACTTGGGCCAGCCAGCCTTGGACAGCACCTTC 990

RESULT 5
AX384610 1773 bp DNA linear PAT 19-MAR-2002
LOCUS
DEFINITION Sequence 7 from Patent WO214503.
ACCESSION AX384610
VERSION AX384610.1 GI:19577811

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1

REFERENCE
AUTHORS McNeill, P.D. and Vedvick, T.S.
TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 7 21-FEB-2002;
CORIXA CORPORATION (US)

FEATURES
source
Location/Qualifiers
1..1773
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 383 a 528 c 530 g 332 t
ORIGIN

Alignment Scores:
Pred. No.: 0.0306 Length: 1773
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AX384610 (1-1773)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 GTGGTCATCCAGAAAGAGGACTTGGGCCAGCCAGCCTTGGACAGCACCTTC 1014

RESULT 6
AX384608 1806 bp DNA linear PAT 19-MAR-2002
LOCUS
DEFINITION Sequence 5 from Patent WO214503.
ACCESSION AX384608
VERSION AX384608.1 GI:19577809

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1

REFERENCE
AUTHORS Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
TITLE Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies

JOURNAL Patent: WO 0214503-A 5 21-FEB-2002;
CORIXA CORPORATION (US)

FEATURES
source
Location/Qualifiers
1..1806
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 391 a 530 c 544 g 341 t
ORIGIN

Alignment Scores:
Pred. No.: 0.0312 Length: 1806
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AX384608 (1-1806)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 994 GTGGTCATCCAGAAAGAGGACTTGGGCCAGCCAGTCTTGGACAGCACCTTC 1047

RESULT 7
AX505114 3678 bp DNA linear PAT 27-SEP-2002
LOCUS
DEFINITION Sequence 7 from Patent WO240059.
ACCESSION AX505114
VERSION AX505114.1 GI:23386421
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1

REFERENCE
AUTHORS Mincheff, M.S., Loukinov, D.I. and Zoubak, S.
TITLE Methods and compositions for inducing cell-mediated immune
responses

JOURNAL Patent: WO 0240059-A 7 23-MAY-2002;
American Foundation for Biological Research Inc. (US) ; Mincheff,
Milcho S. (US) ; Loukinov, Dmitri I. (US) ; Zoubak, Serguei (US)

FEATURES
source
Location/Qualifiers
1..3678
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
CDS
7..3678
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD48804.1"
/db_xref="GI:23386422"

/translation="MRLPASPEHLDMLRHLYVQCQVVOGNLELYLPTNLSLSPLOD
IQEVQYVLIAHNQVQVPLQRLRIVRGTFQEDNYALAVLDNGDPLNNTTPTVGASP
GGLRLQLSLTEILKGLVLTORNPOLCYQDTILWKDIFHKNNOLALTLDITNRSRAC
HPCSPKCGKSGWSSSEDCOSLRTVCAGGCARCKGPLPTCCHEQCAAGCTGPKHS
DCLALHPNHSGICELHCPALVTYNTDTFESMPNPEBGTFCASCATCPYNYLSTDV
GCSLTLCPLHNQEVTAEDGTORCEKCSKPCARVCYGLGNEHLREVRAVTSANIQEFAG
CKKIFGSALFLPESFDGDPASNTAPLQEQVVFETLEITGLYLYISAWPSLPSLV
FQNLQVIRGRLIHNGAYSLTLQGLISWLGRSLRGLSGGLALIHNTHLCFVHTVPW
DQLFRNPHQALLHTANRDEDCVGEGLACHQLCARGHCWGPGTQCVCNCSQFLRGQEC
VEECRVLOGLPREYVNAHRLCPCHPECPQNGSVTCFGEPAOCVACAHYKDPPECA
RCPGSKPDLSPYIWKFPDEBACOPCPINCTHSCVDLDKDGCPAQRASPLNTSIV
AVGILLVVLGVVFGILIKRQOKIRKYTHRRLLQETELVEPLTPSCAMPNQAMRI
LKETELRRKPVKVGSAFTVYKGIWIPDGENVKIPVAKIKRENTSPKANKEILDEAY
VMAGVSPKYSVLLIGLICLTSTVQLTQMLPGLDHRNRRGLSGDGLLNCMQIA
KMSYLEDLRVLRHDLAARNVLKSPNHKITDFGLARLDTDETEYHADGKVPKIKW
MALESILRRRTTHOSDVNSYGVVWELMTFGAKPYDPAEIPDLLLEKGRLPPOPI
CTIDVMTMKVCMIDSCRFRELVSFSESMARDPQRFVVIQNEIDGLASPLDSTF
YRSLEEDDMGDLVDAEYLVFQGFCCPDAPGAGGMVHRHRSSTSSRSGGDDLTIG
LEPSEEEAPRSLAPSEAGSDVDGLMGAAKQLSLPLTHSPLOQYSEDPTVPL
PSETDGYVAPLTCSQPQYVNPQVPQPPQPPREGPLPAARPAAGATLBRAKTLSPGKN
GVWKDVAFGGAIVENPYLTPQGGAAQPHPPPAFSPAFDNLNLYWDQDPPPERGAPPST

BASE COUNT 746 a 1137 c 1091 g 704 t
ORIGIN

FKGTPTAENPEYLGIDVPV"

Alignment Scores:
Pred. No.: 0.0666 Length: 3678
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AX505114 (1-3678)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2869 GTGGTCATCCAGAAATGAGGACTTGGGCCCGACCGAGTCCCTTGGACAGCACCTTC 2922

RESULT 8

AR034479
LOCUS AR034479 3768 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5869445.
ACCESSION AR034479
VERSION AR034479.1 GI:5950084
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3768)

AUTHORS Cheever,M.A. and Disis,M.L.

TITLE Methods for eliciting or enhancing reactivity to HER-2/neu protein

JOURNAL Patent: US 5869445-A 1 09-FEB-1999;

FEATURES Location/Qualifiers

1..3768

source /organism="unknown"

BASE COUNT 759 a 1171 c 1119 g 719 t

ORIGIN

Alignment Scores:
Pred. No.: 0.0683 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AR034479 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGACCGAGTCCCTTGGACAGCACCTTC 3012

RESULT 9

AX060704
LOCUS AX060704 3768 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 2 from Patent WO0100244.
ACCESSION AX060704
VERSION AX060704.1 GI:12406101
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Erickson,S. and Schwall,R.

TITLE Methods of treatment using anti-erbB antibody-maytansinoid

JOURNAL Patent: WO 0100244-A 2 04-JAN-2001;

FEATURES Location/Qualifiers

1..3768

source /organism="Homo sapiens"

/mol_type="genomic DNA"

BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN

/db_xref="taxon:9606"

Alignment Scores:
Pred. No.: 0.0683 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AX060704 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGACCGAGTCCCTTGGACAGCACCTTC 3012

RESULT 10

AX201817
LOCUS AX201817 3768 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0153463.
ACCESSION AX201817
VERSION AX201817.1 GI:15391666
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Cheever,M.A. and Hand-Zimmermann,S.

TITLE Compounds and methods for prevention and treatment of her-2/ neu

JOURNAL Patent: WO 0153463-A 1 26-JUL-2001;

FEATURES Location/Qualifiers

1..3768

source /organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

1..3768

/note="unnamed protein product"

/codon_start=1

/protein_id="C5391667"

/db_xref="GI:15391667"

/translation="MELAAALCRWGLLLALLPGCAASTQVCTGDMKLRLPASPETHLD

MRHLYGCGQVQGNLELYLPFNASISFLQDQVQGYLIAHNQVRQVPLQRLRV

RGQLFEDNTALAVLDNGDPLNNTPTVTGASPGGLRELQRLSLEILKGGVLIQRNPQ

LCYQDTILMKDIFHKNNQALALTIDNTRACHPSPMCKGSCWGESSEDCQSLYRT

VCAGGACRCKGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHSGICELHCPALVYNT

DTPEMPNPGRYTFGACSVTACPNYVLTGVSCTLVCPLNHQVTAEDGTQRCCK

SKPCARVCYGLGMEHLREYRATVTSANIOEFACCKIFGSLAFIPESFDGDPASNTAPL

QPEQLQVFTELEITGYLISAMPDLSLDLVSVFQNLQVIRGRILHNGAYSLTLQGLGI

SWIGLSRLRELGLSLALIHNTLHCFVHTVPDQFRNPHQALLHTANRPEDSCVGGG

LACHQICARHGCHWGPQTQVNCVQSLRGQECVEECVQLQGLPREYVNAHRLCPHPE

CQPNQSVTCFGEADQVACAHYKDPFCVACRCPGVKPDLSYMTWKPPDEGACQ

PCPINTGSCVDLDDKDCPAEORASPLTSTISAVVGLLVVVLGVFGILIKRQOKI

RYTMRLLQETELVEPLTPSGAMPNQAOKRILKTELKVKVLGSGAFGVYKGI

PUGENRKIPVAIKVRENTSPKANKELDEAYVMAGVGSFVSRLLGICLTSTVQVLT

QLMPYGCILLDHRENRGLSGDQLLNCWQIAKMSYLEDLRVLRHLDAAARNLVKSP

NHVTIDFGLRLDIDETEHADGKVPFKWMALESILRRFTHOSVMSYGVYWE

LMTFGAKPDGIPARETIDLEKGERLPQPICTIDVIMIMVCMWIDSCRRPRFREL

VSPFSRMARDPQRFVIONEDLGPASPLDSTFVRSLEDDMDGLDABEYLVPOQGF

FCDDPAPAGGAMVHHRSSSTRSGGDLTLGLEPSEEAAPRSLAPSEGAGSDVDG

DUGMAAKQLSLPHTDPSPLQRYSEDPTVPLPSPTDGYVAPLTCSPPQYIVNPDVR

PQPPSPREGPLPAAPAGATLERPKTLSPKNGVGVKDVAFAGAVENPEYLTPOGGAA

PQPHPPAFSPAFDNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVPV"

BASE COUNT 759 a 1171 c 1119 g 719 t

ORIGIN

Alignment Scores:
Pred. No.: 0.0683 Length: 3768
Score: 83.00 Matches: 16

JUNC_SEQ8_SEQ4 (1-30) x AX384604 (1-3768)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 13
AX465456
LOCUS AX465456 3768 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO0213847.
ACCESSION AX465456
VERSION AX465456.1 GI:21899819
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Gaiger, A., Cheever, M.A. and Hand-Zimmermann, S.
TITLE Methods for diagnosis and therapy of hematological and virus-associated malignancies
JOURNAL Patent: WO 0213847-A 1 21-FEB-2002;
CORIXA CORPORATION (US)

FEATURES
source
Location/Qualifiers
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

CDS
1..3768
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD42363.1"
/db_xref="GI:21899820"

/translation="MELAALCRWGLLLALLPPGAASQVCTGTDMKLRLPASPEHLD
MRHLYOGCVVQGNLELYLPTNASLFLQDIOEVGYVLIHNOVQVPLRLRIV
RGTLQFEDNALAVLDNGDPLNNTPTVPGASGGRLQLRSLTEILKGGVLIQRNPQ
LCYQDTILMKDIFHKNOLALTLIDTNRACHPSPCKSGRSGWSSDCCSLRT
VCAGGACRCKGPLETDCHEQACAGCTGPKHSDCLACLFHNSGICELHCPALVTYNT
DTFSPMPNPGRYTFGASCTACPYNTLSDVGSCTLVCPHNOEVAEDGTQRCCK
SKPCARVYGLGMEHLEVRVAVTSANTQEFAGCKIFGSLAFLPESDGDPSANTPL
OPEQLOVFEETLEETGYLYISAWPDSLPLDSVFONLOVIRGRILHNGAYSLTLOGLI
SWGLSLRELGSGLALIHNTLHLCFVHTVPDQLFRNPHQALLHTANRDECEVGG
LACHQLCARGHCGPGTQCVCNCSQFLRGQECVEECVQLGPREYVNRHCLPCHPE
PCPNCTHSCVDLDKGCPCAEQASPLTSISAVGILLVVLVGVFGILIKRQKI
RKYTMRLLOETELVEPLTPSGAMPNQAKRILKTELKRVKVLGSGAFVTKGIWI
PDGNNVPAIKVLRNTSPKANKEILDEAYVMAGVSPVSVRLGICLTSTVOLVT
QLMPYGCILLDHVNRGRIGSODLLNMCQIAKMSYLEDVRLVHRLAARNVLYKSP
NHVKITDFGLARLLDIDETEHADGGKVPINWMALESILRRRFTHQSDVWSYGVYWE
LMTFGAPYDGPAREIPDLLEKGERLPQPTCTIDVIMVKNWCMIDSECRPRFREL
VSEFSRMARDPQRFVVIQNEDLGSPALDSTFYRSLLEDMDGDLVDAEYLVPOQFG
FCPDPAAGAGVHRRHSSTRSGGDLTLGLEPSEEEAPRSLAPSEGAAGSVDFG
DLGMAAGKQLSLTPHDSPLQRYSEDPTVLPSETDGYVAPLTCSPQPEVYNQPDVR
PQPPSPREGPLPAARPAAGATLERKTLSPCKNGVYKDVFAFGGAVENPEYLTPOGGAA
PQHPHPPAFSAFDNLYYWDODPBERGAPSTFKGTPTAENPEYLGLDVVP"

BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN

Alignment Scores:
Pred. No.: 0.0683 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AX465456 (1-3768)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 14

AX467229
LOCUS AX467229 3768 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0234287.
ACCESSION AX467229
VERSION AX467229.1 GI:21900511
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Beier, A.M., Gautam, A. and Mouritsen, S.R.
TITLE Novel therapeutic vaccine formulations
JOURNAL Patent: WO 0234287-A 3 02-MAY-2002;
Pharmexa A/S (DK)

FEATURES
source
Location/Qualifiers
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

CDS
1..3768
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD42452.1"
/db_xref="GI:21900512"

/translation="MELAALCRWGLLLALLPPGAASQVCTGTDMKLRLPASPEHLD
MRHLYOGCVVQGNLELYLPTNASLFLQDIOEVGYVLIHNOVQVPLRLRIV
RGTLQFEDNALAVLDNGDPLNNTPTVPGASGGRLQLRSLTEILKGGVLIQRNPQ
LCYQDTILMKDIFHKNOLALTLIDTNRACHPSPCKSGRSGWSSDCCSLRT
VCAGGACRCKGPLETDCHEQACAGCTGPKHSDCLACLFHNSGICELHCPALVTYNT
DTFSPMPNPGRYTFGASCTACPYNTLSDVGSCTLVCPHNOEVAEDGTQRCCK
SKPCARVYGLGMEHLEVRVAVTSANTQEFAGCKIFGSLAFLPESDGDPSANTPL
OPEQLOVFEETLEETGYLYISAWPDSLPLDSVFONLOVIRGRILHNGAYSLTLOGLI
SWGLSLRELGSGLALIHNTLHLCFVHTVPDQLFRNPHQALLHTANRDECEVGG
LACHQLCARGHCGPGTQCVCNCSQFLRGQECVEECVQLGPREYVNRHCLPCHPE
PCPNCTHSCVDLDKGCPCAEQASPLTSISAVGILLVVLVGVFGILIKRQKI
RKYTMRLLOETELVEPLTPSGAMPNQAKRILKTELKRVKVLGSGAFVTKGIWI
PDGNNVPAIKVLRNTSPKANKEILDEAYVMAGVSPVSVRLGICLTSTVOLVT
QLMPYGCILLDHVNRGRIGSODLLNMCQIAKMSYLEDVRLVHRLAARNVLYKSP
NHVKITDFGLARLLDIDETEHADGGKVPINWMALESILRRRFTHQSDVWSYGVYWE
LMTFGAPYDGPAREIPDLLEKGERLPQPTCTIDVIMVKNWCMIDSECRPRFREL
VSEFSRMARDPQRFVVIQNEDLGSPALDSTFYRSLLEDMDGDLVDAEYLVPOQFG
FCPDPAAGAGVHRRHSSTRSGGDLTLGLEPSEEEAPRSLAPSEGAAGSVDFG
DLGMAAGKQLSLTPHDSPLQRYSEDPTVLPSETDGYVAPLTCSPQPEVYNQPDVR
PQPPSPREGPLPAARPAAGATLERKTLSPCKNGVYKDVFAFGGAVENPEYLTPOGGAA
PQHPHPPAFSAFDNLYYWDODPBERGAPSTFKGTPTAENPEYLGLDVVP"

BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN

Alignment Scores:
Pred. No.: 0.0683 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AX467229 (1-3768)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 15
AX481438
LOCUS AX481438 3768 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 52 from Patent WO02055693.
ACCESSION AX481438
VERSION AX481438.1 GI:22316352
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.
Method for inhibiting the expression of a target gene
Patent: WO 02055693-A 52 18-JUL-2002;
Ribopharma AG (DE)

FEATURES

Location/Qualifiers
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN

Alignment Scores:

Pred. No.:	0.0683	Length:	3768
Score:	83.00	Matches:	16
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	2
Query Match:	51.88%	Indels:	0
DB:	6	Gaps:	0

JUNC_SEQ8_SEQ4 (1-30) x AX481438 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
||| |||||
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGCCAGCTCCCTTGGACAGCACCTTC 3012

Search completed: October 16, 2003, 11:01:46
Job time : 436.208 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 234.794 seconds
(without alignments)
3105.423 Million cell updates/sec

Title: JUNC_SEQ8_SEQ2RES991_

Perfect score: 161

Sequence: 1 ERGCPAEQASQPTVFNEDLGSPSPMDSTF 30

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delep 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool/P/OLLERAN480/runat_15102003_131913_20548/app_query.fasta_1.4685
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 @CGN_1_1_12645 @runat_15102003_131913_20548 -NCPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vit:*

21: em_gss_fun:*

22: em_gss_man:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vri:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	84	52.2	422	10	BF554233	BF554233 UI-R-C0-h
2	84	52.2	524	13	BQ828479	BQ828479 LL6in1002
3	84	52.2	527	14	CB719751	CB719751 AMGNNUC:N
4	84	52.2	568	13	BQ831943	BQ831943 LL6in1122
5	84	52.2	577	13	BQ829915	BQ829915 LL6in2239
6	84	52.2	592	10	BE382065	BE382065 601272225
7	84	52.2	629	10	BE374596	BE374596 601225384
8	84	52.2	666	13	BU671195	BU671195 NISC_lr06
9	84	52.2	715	13	BU052728	BU052728 UI-M-EX0-
10	84	52.2	726	12	BI648082	BI648082 603279290
11	84	52.2	770	12	BI645946	BI645946 603274921
12	84	52.2	773	12	BI157032	BI157032 602921306
13	84	52.2	778	12	BG863431	BG863431 602979093
14	84	52.2	785	12	BI688650	BI688650 603311262
15	84	52.2	787	12	BI156730	BI156730 602922150
16	84	52.2	844	12	BI158626	BI158626 602922395
17	84	52.2	881	12	BI645770	BI645770 603275203
18	84	52.2	887	14	CA980253	CA980253 AGENCOURT
19	84	52.2	984	13	BU523863	BU523863 AGENCOURT
20	84	52.2	1100	11	BC023725	BC023725 Mus muscu
21	84	52.2	1110	11	AK031542	AK031542 Mus muscu
22	84	52.2	3372	11	BC046553	BC046553 Mus muscu
23	84	52.2	4323	11	AK031099	AK031099 Mus muscu
24	84	52.2	4463	11	AK083669	AK083669 Mus muscu
25	81.5	50.6	618	28	AZ699313	AZ699313 RPECI-23-2
26	81	50.3	220	9	AW763060	AW763060 ur60c11.y
27	81	50.3	483	4	BX530001	BX530001 RZPD Mus
28	81	50.3	552	10	BF022141	BF022141 uv47c11.y
29	81	50.3	654	12	BI648646	BI648646 603277119
30	81	50.3	655	12	BI557778	BI557778 603236957
31	81	50.3	695	12	BI1557170	BI1557170 603239144
32	81	50.3	718	12	BI108229	BI108229 602902263
33	81	50.3	749	12	BI554781	BI554781 603236758
34	81	50.3	754	12	BI651822	BI651822 603298821
35	81	50.3	771	12	BI154609	BI154609 602904778
36	81	50.3	891	10	BF101641	BF101641 601753564
37	78	48.4	278	10	BF923905	BF923905 QV4-NT025
38	78	48.4	294	10	BF757857	BF757857 CM4-CT057
39	78	48.4	337	10	BE091682	BE091682 IL2-BT073
40	78	48.4	348	9	AI909847	AI909847 QV-BT225-
41	78	48.4	378	14	CB2666492	CB2666492 1005398 H
42	78	48.4	423	12	BM703962	BM703962 UI-E-CK1-
43	78	48.4	464	2	HSM067221	Bx477782 Homo sapi
44	78	48.4	496	12	BM795191	BM795191 K-EST0076
45	78	48.4	505	12	BM854421	BM854421 K-EST0136

ALIGNMENTS

RESULT 1
BF554233

LOCUS
DEFINITION

BF554233 422 bp mRNA linear EST 12-DEC-2000
UI-R-C0-hp-c-01-0-UI.r1 UI-R-C0 Rattus norvegicus cDNA clone
UI-R-C0-hp-c-01-0-UI 5', mRNA sequence.

ACCESSION
VERSION

BF554233
BF554233.1 GI:11663963

KEYWORDS
EST.

SOURCE
ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 422)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@iowa.edu
 cdna Library Preparation: M.B. Soares Lab Clone distribution:
 clones will be available through Research Genetics (www.resgen.com)
 This clone is also available through the I.M.A.G.E. Consortium at
 LLNL (info@image.llnl.gov). IMAGE ID= 1781731
 Seq primer: M13 Forward.

FEATURES

Location/Qualifiers
 1..422
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C0-hp-c-01-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-C0"
 /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; The UI-R-C0 library is a subtracted library derived from the UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, and muscle. The UI-R-E1 library consisted of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C0) was constructed as follows: PCR amplified cDNA inserts from a pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the pooled UI-R-A1 and UI-R-E1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C0 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)."

BASE COUNT 90 a 115 c 125 g 92 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0.0938 Length: 422
 Score: 84.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 52.17% Indels: 0
 DB: 10 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x BF554233 (1-422)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
 ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 122 GTGGTCATCCAGAACGAGGACTTGGGCCCATCCAGCCCATGGACAGTACCTTC 175
 RESULT 2
 BQ828479

LOCUS BQ828479 524 bp mRNA linear EST 15-SEP-2002
 DEFINITION LL6in1002577 AFT024-subtracted library Mus musculus cDNA 5' similar to NeuYERBB-2 p185, mRNA sequence.
 ACCESSION BQ828479
 VERSION BQ828479.1 GI:22860534
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 524)
 REFERENCE Hackney,J.A., Charbord,P., Brunk,B.P., Stoeckert,C.J., Lemischka,I.R. and Moore,K.A.
 A molecular profile of a hematopoietic stem cell niche
 Proc. Natl. Acad. Sci. U.S.A. 99 (20), 13061-13066 (2002)
 JOURNAL 22247628
 MEDLINE 12226475
 PUBMED
 COMMENT Contact: Moore, Kateri A.
 Department of Molecular Biology
 Princeton University
 217 Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544, USA
 Tel: 609 258 0605
 Fax: 609 258 2759
 Email: kmoore@molbio.princeton.edu
 These ESTs are derived from a subtracted cDNA library enriched for gene products expressed by a hematopoietic stem cell-supporting stromal cell line, AFT024.
 Seq primer: M13Reverse or T7.

FEATURES

Location/Qualifiers
 1..524
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /tissue_type="Fetal Liver"
 /cell_type="Stromal cell"
 /cell_line="AFT024"
 /dev_stage="Embryonic day 14-14.5"
 /lab_host="DH10B"
 /clone_lib="AFT024-subtracted library"
 /notes="Organ: Fetal Liver; Vector: Sport 1; Site.1: Sal I; Site.2: Not I; Two directionally cloned cDNA libraries were made from a hematopoietic stem cell-supporting stromal cell line (AFT024) and from a non-supporting stromal cell line (2018). Subtractive hybridization was performed by hybridization of the target, AFT024, single stranded cDNA library in psport1 to biotinylated RNA transcribed from the driver, 2018 cDNA library in pSport2 with inserts cloned in the complementary orientation. The AFT024-subtracted library contains 4.2x105 clones and is depleted of common housekeeping gene products eg. beta-actin and enriched for transcripts specific to AFT024. For detailed protocols and additional information please see our website at
 http://stromalcell.princeton.edu."

BASE COUNT 110 a 149 c 157 g 107 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0.124 Length: 524
 Score: 84.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 52.17% Indels: 0
 DB: 13 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x BQ828479 (1-524)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
 ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 270 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCATGGACAGTACCTTC 323

RESULT 3

CB719751
 LOCUS
 DEFINITION AMGNNUC:NRDGI-00186-C7-A nrdgl (10855) Rattus norvegicus cDNA clone
 nrdgl-00186-c7 5', mRNA sequence.

ACCESSION CB719751

VERSION CB719751

KEYWORDS CB719751.1 GI:29776893

SOURCE EST.

ORGANISM Rattus norvegicus (Norway rat)

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 527)

AUTHORS Angen EST Program

TITLE Angen Rat EST Program

JOURNAL Unpublished

COMMENT Contact: Dan Fitzpatrick

Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00186 row: c column: 7.

Location/Qualifiers

1..527

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="nrdgl-00186-c7"

/tissue_type="Dorsal Root Ganglia"

/clone_lib="nrdgl (10855)"

/note="vector: pSPORT1; Site_1: SalI; Site_2: NotI; rat
 dorsal root ganglia"

BASE COUNT 105 a 159 c 156 g 107 t

ORIGIN

Alignment Scores:
 Pred. NO.: 0.125 Length: 527
 Score: 84.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 52.17% Indels: 0
 DB: 14 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x CB719751 (1-527)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30

Db 87 GTGTCATCCAGAACGAGGACTTGGCCCATCCAGCCCATGGACAGTACCTTC 140

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

Tel: 609 258 0605

Fax: 609 258 2759

Email: knoore@molbio.princeton.edu

These ESTs are derived from a subtracted cDNA library enriched for
 gene products expressed by a hematopoietic stem cell-supporting
 stromal cell line, AFT024.

Seq primer: M13Reverse or T7.

Location/Qualifiers

1..568

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57Bl/6J"

/db_xref="taxon:10090"

/tissue_type="Fetal Liver"

/cell_type="Stromal cell"

/cell_line="AFT024"

/dev_stage="Embryonic day 14-14.5"

/lab_host="DH10B"

/clone_lib="AFT024-subtracted library"

/note="Organ: Fetal Liver; Vector: Sport 1; Site_1: Sal I;

Site_2: Not I; Two directionally cloned cDNA libraries

were made from a hematopoietic stem cell-supporting

stromal cell line (AFT024) and from a non-supporting

stromal cell line (2018). Subtractive hybridization was

performed by hybridization of the target, AFT024, single

stranded cDNA library in pSport1 to biotinylated RNA

transcribed from the driver, 2018 cDNA library in pSport2

with inserts cloned in the complementary orientation. The

AFT024-subtracted library contains 4.2x10⁵ clones and is

depleted of common housekeeping gene products eg.

beta-actin and enriched for transcripts specific to

AFT024. For detailed protocols and additional information

please see our website at

http://stromalcell.princeton.edu."

BASE COUNT 122 a 173 c 159 g 112 t 2 others

ORIGIN

Alignment Scores:
 Pred. NO.: 0.138 Length: 568
 Score: 84.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 52.17% Indels: 0
 DB: 13 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x B0831943 (1-568)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30

Db 156 GTGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCATGGACAGCCTTC 209

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

RESULT 4

CB719751
 LOCUS
 DEFINITION AMGNNUC:NRDGI-00186-C7-A nrdgl (10855) Rattus norvegicus cDNA clone
 nrdgl-00186-c7 5', mRNA sequence.

ACCESSION CB719751

VERSION CB719751

KEYWORDS CB719751.1 GI:29776893

SOURCE EST.

ORGANISM Rattus norvegicus (Norway rat)

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 527)

AUTHORS Angen EST Program

TITLE Angen Rat EST Program

JOURNAL Unpublished

COMMENT Contact: Dan Fitzpatrick

Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00186 row: c column: 7.

Location/Qualifiers

1..527

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="nrdgl-00186-c7"

/tissue_type="Dorsal Root Ganglia"

/clone_lib="nrdgl (10855)"

/note="vector: pSPORT1; Site_1: SalI; Site_2: NotI; rat
 dorsal root ganglia"

BASE COUNT 105 a 159 c 156 g 107 t

ORIGIN

Alignment Scores:
 Pred. NO.: 0.125 Length: 527
 Score: 84.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 52.17% Indels: 0
 DB: 14 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x CB719751 (1-527)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30

Db 87 GTGTCATCCAGAACGAGGACTTGGCCCATCCAGCCCATGGACAGTACCTTC 140

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

RESULT 5

CB719751
 LOCUS
 DEFINITION AMGNNUC:NRDGI-00186-C7-A nrdgl (10855) Rattus norvegicus cDNA clone
 nrdgl-00186-c7 5', mRNA sequence.

ACCESSION CB719751

VERSION CB719751

KEYWORDS CB719751.1 GI:29776893

SOURCE EST.

ORGANISM Rattus norvegicus (Norway rat)

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 527)

AUTHORS Angen EST Program

TITLE Angen Rat EST Program

JOURNAL Unpublished

COMMENT Contact: Dan Fitzpatrick

Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00186 row: c column: 7.

Location/Qualifiers

1..527

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="nrdgl-00186-c7"

/tissue_type="Dorsal Root Ganglia"

/clone_lib="nrdgl (10855)"

/note="vector: pSPORT1; Site_1: SalI; Site_2: NotI; rat
 dorsal root ganglia"

BASE COUNT 105 a 159 c 156 g 107 t

ORIGIN

Alignment Scores:
 Pred. NO.: 0.125 Length: 527
 Score: 84.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 52.17% Indels: 0
 DB: 14 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x CB719751 (1-527)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30

Db 87 GTGTCATCCAGAACGAGGACTTGGCCCATCCAGCCCATGGACAGTACCTTC 140

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

111 |||||

217 Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544,
USA
Tel: 609 258 0605
Fax: 609 258 2759
Email: kmoore@molbio.princeton.edu
These ESTs are derived from a subtracted cDNA library enriched for
gene products expressed by a hematopoietic stem cell-supporting
stromal cell line, AFT024.
Seq primer: M13reverse or T7.

FEATURES

source
1. .577
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/tissue_type="Fetal Liver"
/cell_type="Stromal cell"
/cell_line="AFT024"
/dev_stage="Embryonic day 14-14.5"
/lab_host="DH10B"
/clone_lib="AFT024-subtracted library"
/notes="Organ: Fetal Liver; Vector: Sport 1; Site_1: Sal I;
Site_2: Not I; Two directionally cloned cDNA libraries
were made from a hematopoietic stem cell-supporting
stromal cell line (AFT024) and from a non-supporting
stromal cell line (2018). Subtractive hybridization was
performed by hybridization of the target, AFT024, single
stranded cDNA library in pSport1 to biotinylated RNA
transcribed from the driver, 2018 cDNA library in pSport2
with inserts cloned in the complementary orientation. The
AFT024-subtracted library contains 4.2x105 clones and is
depleted of common housekeeping gene products eg.
beta-actin and enriched for transcripts specific to
AFT024. For detailed protocols and additional information
please see our website at
http://stromalcell.princeton.edu."

BASE COUNT 123 a 173 c 170 g 110 t
ORIGIN
Alignment Scores:
Pred. No.: 0.141 Length: 577
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 13 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x B0829915 (1-577)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 167 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCATGGACGACCTTC 220

RESULT 6
BE382065 592 bp mRNA linear EST 21-JUL-2000
LOCUS 601272225F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3601422 5',
DEFINITION mRNA sequence.
ACCESSION BE382065.1 GI:9327430
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 592)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.

JUNC_SEQ8_SEQ2RES991_ (1-30) x B0829915 (1-577)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 167 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCATGGACGACCTTC 220

RESULT 6
BE382065 592 bp mRNA linear EST 21-JUL-2000
LOCUS 601272225F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3601422 5',
DEFINITION mRNA sequence.
ACCESSION BE382065.1 GI:9327430
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 592)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.

FEATURES
source
1. .577
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/tissue_type="Fetal Liver"
/cell_type="Stromal cell"
/cell_line="AFT024"
/dev_stage="Embryonic day 14-14.5"
/lab_host="DH10B"
/clone_lib="AFT024-subtracted library"
/notes="Organ: Fetal Liver; Vector: Sport 1; Site_1: Sal I;
Site_2: Not I; Two directionally cloned cDNA libraries
were made from a hematopoietic stem cell-supporting
stromal cell line (AFT024) and from a non-supporting
stromal cell line (2018). Subtractive hybridization was
performed by hybridization of the target, AFT024, single
stranded cDNA library in pSport1 to biotinylated RNA
transcribed from the driver, 2018 cDNA library in pSport2
with inserts cloned in the complementary orientation. The
AFT024-subtracted library contains 4.2x105 clones and is
depleted of common housekeeping gene products eg.
beta-actin and enriched for transcripts specific to
AFT024. For detailed protocols and additional information
please see our website at
http://stromalcell.princeton.edu."

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8787 row: a column: 07
High quality sequence stop: 575.

FEATURES

source
1. .592
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="IMAGE:3601422"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam1"
/notes="Organ: mammary; Vector: pCMV-Sport6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 116 a 194 c 169 g 113 t
ORIGIN

Alignment Scores:
Pred. No.: 0.145 Length: 592
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x BE382065 (1-592)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 76 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCATGGACGACCTTC 129

RESULT 7
BE374596

LOCUS 601225384F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3583556 5',
DEFINITION mRNA sequence.
ACCESSION BE374596
VERSION BE374596.1 GI:9319959
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 629)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8740 row: h column: 21
High quality sequence stop: 580.

FEATURES
source
1. .629
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="IMAGE:3583556"

/tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Mam1"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 118 a 204 c 181 g 126 t

ORIGIN

Alignment Scores:

Pred. No.: 0.157 Length: 629
 Score: 84.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 52.17% Indels: 0
 DB: 10 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x BE374596 (1-629)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerPrometAspSerThrphe 30

Db 8 GTGGTCATCCAGACGAGGACTTAGCCCTCCAGCCCATGGACAGCACCTTC 61

RESULT 8

BU671195

LOCUS

DEFINITION NISC_Lr06b08.y1 NCI_CGAP_Pr49 Rattus norvegicus cDNA clone
 IMAGE:5598279 5', mRNA sequence.

ACCESSION

BU671195

VERSION

BU671195.1 GI:23399536

KEYWORDS

EST.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

1 (bases 1 to 666)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

Unpublished

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

CDNA Library Preparation:
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL

DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC)

Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: L1AM12385 Row: C Column: 16

Seq primer: M13Rpl reverse primer (ABI).

Location/Qualifiers

1..666

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="IMAGE:5598279"

/sex="male"

/tissue_type="ventral prostate, pool of 3-, 5-, and 7-days

post-castration"

/dev_stage="adult, 11 week"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NCI_CGAP_Pr49"

/note="Organ: prostate; Vector: pCMV-SPORT6.cdb; Site_1:
 NotI; Site_2: EcoRV; Cloned unidirectionally. Primer:
 Oligo dt. Pool of 3 primary libraries: NCI_CGAP_Pr30
 (ventral prostate from 11 wk male, 3 days post-castration
 , average insert size 2 kb), NCI_CGAP_Pr40 (ventral
 prostate from 11 wk male, 5 days post-castration, average
 insert size 1.6 kb) and NCI_CGAP_Pr41 (ventral prostate

from 11 wk male, 7 days post-castration, average insert
 size 2.5 kb). Constructed by Life
 Technologies/Invitrogen. Note: this is a NCI_CGAP
 Library."

BASE COUNT 149 a 165 c 197 g 155 t

ORIGIN

Alignment Scores:

Pred. No.: 0.17 Length: 666
 Score: 84.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 52.17% Indels: 0
 DB: 13 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x BU671195 (1-666)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerPrometAspSerThrphe 30

Db 563 GTGGTCATCCAGACGAGGACTTGGCCATCCAGCCCATGGACAGTACCTTC 616

RESULT 9

BU052728

LOCUS

DEFINITION UI-M-EX0-b-yk-b-08-0-UI.r1 NIH_BMAP_EX0 Mus musculus cDNA clone
 IMAGE:5719207 5', mRNA sequence.

ACCESSION

BU052728

VERSION

BU052728.1 GI:22492805

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 715)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..715

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5719207"

/tissue_type="whole brain"

/dev_stage="embryo 15.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_EX0"

/note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 , is GTGGCTGGAA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous

FEATURES
source

Location/Qualifiers
1. .785
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5347334"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 156 a 265 c 214 g 150 t
ORIGIN

Alignment Scores:
Pred. No.: 0.21 Length: 785
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x BI688650 (1-785)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 100 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCATGGACAGCACCTTC 153

RESULT 15

BI156730
LOCUS 787 bp mRNA linear EST 05-JUL-2001
DEFINITION 602922150F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5062237 5',
mRNA sequence.
ACCESSION BI156730
VERSION BI156730.1 GI:14616731
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 787)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1168 row: d column: 14
High quality sequence stop: 748.

FEATURES
source

Location/Qualifiers
1. .787
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129/C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5062237"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."

BASE COUNT 166 a 253 c 214 g 154 t
ORIGIN

Alignment Scores:
Pred. No.: 0.211 Length: 787
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x BI156730 (1-787)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 68 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCATGGACAGCACCTTC 121

Search completed: October 16, 2003, 17:04:04
Job time : 235.794 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: October 16, 2003, 11:08:07 : Search time 31.0111 Seconds
(without alignments)
2540.503 Million cell updates/sec

Title: JUNC_SEQ8_SEQ2RES991_
Perfect score: 161
Sequence: 1 ERGCPAQRASPVTFQNEGLGPPSPMDSTF 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPto_spool_p/HOLLERAN480/runat_15102003_131915_20662/app_query.fasta_1.4685
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NAXLEN=2000000000 -USER=HOLLERAN480 -SCGN_1_1_397 @runat_15102003_131915_20662
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	84	52.2	3771	10 US-09-854-356-11	Sequence 11, Appl

2	84	52.2	3955	10	US-09-870-759-117	Sequence 117, App
3	84	52.2	3955	10	US-09-854-356-10	Sequence 10, Appl
4	84	52.2	3955	12	US-09-751-708A-117	Sequence 117, App
5	78	48.4	1115	14	US-10-102-806-165	Sequence 165, App
6	78	48.4	1713	12	US-09-378-393-14	Sequence 14, Appl
7	78	48.4	1755	10	US-09-930-125-6	Sequence 6, Appl
8	78	48.4	1767	10	US-09-930-125-7	Sequence 7, Appl
9	78	48.4	1773	10	US-09-930-125-4	Sequence 4, Appl
10	78	48.4	1806	10	US-09-930-125-5	Sequence 5, Appl
11	78	48.4	2411	12	US-10-378-393-10	Sequence 10, Appl
12	78	48.4	3765	12	US-10-207-498-5	Sequence 5, Appl
13	78	48.4	3768	9	US-09-811-123-8	Sequence 8, Appl
14	78	48.4	3768	9	US-09-811-115-2	Sequence 2, Appl
15	78	48.4	3768	10	US-09-854-356-9	Sequence 9, Appl
16	78	48.4	3768	10	US-09-930-125-1	Sequence 1, Appl
17	78	48.4	3768	12	US-10-313-644-1	Sequence 1, Appl
18	78	48.4	4473	11	US-09-441-411-5	Sequence 5, Appl
19	78	48.4	4473	12	US-10-101-510-81	Sequence 81, Appl
20	78	48.4	4473	14	US-10-146-473-32	Sequence 32, Appl
21	78	48.4	4473	14	US-10-207-655-44	Sequence 44, Appl
22	78	48.4	4530	10	US-09-877-177-11	Sequence 11, Appl
23	78	48.4	4530	12	US-10-007-526A-119	Sequence 119, App
24	78	48.4	4530	12	US-10-101-510-124	Sequence 124, App
25	78	48.4	4530	12	US-10-338-730-1	Sequence 1, Appl
26	78	48.4	4530	14	US-10-177-293-125	Sequence 125, App
27	78	48.4	4543	10	US-09-769-508-1	Sequence 1, Appl
28	78	48.4	4606	12	US-09-971-392-70	Sequence 70, Appl
29	78	48.4	4642	14	US-10-198-846-10896	Sequence 10896, A
30	78	48.4	9274	9	US-09-811-123-7	Sequence 7, Appl
31	78	48.4	9274	9	US-09-811-115-1	Sequence 1, Appl
32	66	41.0	201	13	US-10-109-213-3	Sequence 3, Appl
33	60	37.3	23432	9	US-09-764-869-1332	Sequence 1332, Ap
34	60	37.3	23432	14	US-10-091-504-1332	Sequence 25, Appl
35	59	36.6	1938	12	US-10-422-264-25	Sequence 27, Appl
36	59	36.6	2316	12	US-10-422-264-27	Sequence 23, Appl
37	59	36.6	2604	12	US-10-422-264-23	Sequence 1, Appl
38	59	36.6	3320	10	US-09-862-658-1	Sequence 22, Appl
39	59	36.6	3320	14	US-10-175-696-22	Sequence 22, Appl
40	59	36.6	3384	12	US-10-422-264-29	Sequence 29, Appl
41	57.5	35.7	363	11	US-09-318-995-28706	Sequence 28706, A
42	57.5	35.7	2691	10	US-09-967-768A-292	Sequence 292, App
43	57.5	35.7	2691	13	US-10-028-748-1	Sequence 1, Appl
44	57.5	35.7	2785	13	US-10-002-600-66	Sequence 66, Appl
45	57	35.4	602	13	US-10-027-632-114433	Sequence 114433,

ALIGNMENTS

RESULT 1
US-09-854-356-11
; Sequence 11, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 3771
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse HER-2/neu

US-09-854-356-11

Alignment Scores:

Pred. No.: 0.00383 Length: 3771
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-854-356-11 (1-3771)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30

Db 2962 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCATGCAGACGACACCTTC 3015

RESULT 2

US-09-870-759-117

; Sequence 117, Application US/09870759

; Patent No. US20020177551A1

; GENERAL INFORMATION:

; APPLICANT: Terman, David S

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE

; FILE REFERENCE: 870759

; CURRENT APPLICATION NUMBER: US/09/870,759

; PRIOR FILING DATE: 2002-01-14

; PRIOR APPLICATION NUMBER: US 60/208,128

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 117

; LENGTH: 3955

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (17)..(3799)

; OTHER INFORMATION:

US-09-870-759-117

Alignment Scores:

Pred. No.: 0.00403 Length: 3955
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-870-759-117 (1-3955)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30

Db 2990 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCATGCAGACGACGACCTTC 3043

RESULT 3

US-09-854-356-10

; Sequence 10, Application US/09854356

; Patent No. US20020177567A1

; GENERAL INFORMATION:

; APPLICANT: Cheever, Martin A.

; APPLICANT: Gheysen, Dirk

; APPLICANT: Corixa Corporation

; APPLICANT: Smithline Beecham Biologicals S. A.

; TITLE OF INVENTION: HER-2/neu Fusion Proteins

; FILE REFERENCE: 014058-009810PC

; CURRENT APPLICATION NUMBER: US/09/854,356

; CURRENT FILING DATE: 2001-05-09

; PRIOR APPLICATION NUMBER: US 09/493,480

; PRIOR FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: US 60/117,976

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 10
; LENGTH: 3955
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: rat HER-2/neu cDNA
; NAME/KEY: CDS
; LOCATION: (26)..(3799)
; OTHER INFORMATION: rat HER-2/neu protein
; NAME/KEY: misc.feature
; LOCATION: (26)..(1990)
; OTHER INFORMATION: extracellular domain (ECD) of rat HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2057)..(3796)
; OTHER INFORMATION: intracellular domain (ICD) of rat HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2188)..(3022)
; OTHER INFORMATION: kinase domain (KD) of rat HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2999)..(3796)
; OTHER INFORMATION: phosphorylation domain (PD) of rat HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2999)..(3173)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of rat HER-2/neu
US-09-854-356-10

Alignment Scores:

Pred. No.: 0.00403 Length: 3955
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-854-356-10 (1-3955)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30

Db 2990 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCATGCAGACGACGACCTTC 3043

RESULT 4

US-09-751-708A-117

; Sequence 117, Application US/09751708A

; Publication No. US20030157113A1

; GENERAL INFORMATION:

; APPLICANT: Terman, David S

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE

; FILE REFERENCE: 751708

; CURRENT APPLICATION NUMBER: US/09/751,708A

; CURRENT FILING DATE: 2002-10-15

; PRIOR APPLICATION NUMBER: US 60/173,371

; PRIOR FILING DATE: 1999-12-28

; NUMBER OF SEQ ID NOS: 166

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 117

; LENGTH: 3955

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (17)..(3799)

; OTHER INFORMATION:

US-09-751-708A-117

Alignment Scores:

Pred. No.: 0.00403 Length: 3955
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-751-708A-117 (1-3955)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerPrometAspSerThrPhe 30
Db 2990 GTGGTCATCCAGAAAGAGGACTTGGGCCCATCCAGCCCGGACGAGTACCTTC 3043

RESULT 5

US-10-102-806-165
; Sequence 165, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL03P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 165
; LENGTH: 1115
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (390)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-165

Alignment Scores:
Pred. No.: 0.0085 Length: 1115
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 14 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-10-102-806-165 (1-1115)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerPrometAspSerThrPhe 30
Db 446 GTGGTCATCCAGAAATGAGGACTTGGGCCCGGACGAGTCCCTTGGACAGCACCTTC 499

RESULT 6

US-10-378-393-14
; Sequence 14, Application US/10378393
; Publication No. US20030182668A1
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1713
; TYPE: DNA

; ORGANISM: Homo sapiens
US-10-378-393-14

Alignment Scores:
Pred. No.: 0.0136 Length: 1713
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-10-378-393-14 (1-1713)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerPrometAspSerThrPhe 30
Db 904 GTGGTCATCCAGAAATGAGGACTTGGGCCCGGACGAGTCCCTTGGACAGCACCTTC 957

RESULT 7

US-09-930-125-6
; Sequence 6, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-6

Alignment Scores:
Pred. No.: 0.014 Length: 1755
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-930-125-6 (1-1755)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerPrometAspSerThrPhe 30
Db 937 GTGGTCATCCAGAAATGAGGACTTGGGCCCGGACGAGTCCCTTGGACAGCACCTTC 990

RESULT 8

US-09-930-125-4
; Sequence 4, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14

```
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-4
Alignment Scores:
Pred. No.: 0.0141 Length: 1767
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 10 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-930-125-4 (1-1767)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 937 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 990
RESULT 9
US-09-930-125-7
; Sequence 7, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-7
Alignment Scores:
Pred. No.: 0.0141 Length: 1773
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 10 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-930-125-7 (1-1773)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 961 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1014
RESULT 10
US-09-930-125-5
; Sequence 5, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-5
Alignment Scores:
Pred. No.: 0.0144 Length: 1806
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 10 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-930-125-5 (1-1806)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 994 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1047
RESULT 11
US-10-378-393-10
; Sequence 10, Application US/10378393
; Publication No. US20030182668A1
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVAT
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-10
Alignment Scores:
Pred. No.: 0.0198 Length: 2411
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 12 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-10-378-393-10 (1-2411)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 1602 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1655
RESULT 12
US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
```

```
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREGULIN AND HER3
; FILE REFERENCE: 30448.103-US-01
; CURRENT APPLICATION NUMBER: US/10/207,498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5

Alignment Scores:
Pred. No.: 0.0323      Length: 3765
Score: 78.00          Matches: 14
Percent Similarity: 88.89%      Conservative: 2
Best Local Similarity: 77.78%    Mismatches: 2
Query Match: 48.45%             Indels: 0
DB: 12                      Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-10-207-498-5 (1-3765)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTTGGACAGCACCTTC 3012

RESULT 13
US-09-811-123-8
; Sequence 8, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; TITLE OF INVENTION: ANTIBODY-MAYTANSINOID CONJUGATES
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-123-8

Alignment Scores:
Pred. No.: 0.0323      Length: 3768
Score: 78.00          Matches: 14
Percent Similarity: 88.89%      Conservative: 2
Best Local Similarity: 77.78%    Mismatches: 2
Query Match: 48.45%             Indels: 0
DB: 9                      Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-811-123-8 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTTGGACAGCACCTTC 3012
```

```
RESULT 14
US-09-811-115-2
; Sequence 2, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred. No.: 0.0323      Length: 3768
Score: 78.00          Matches: 14
Percent Similarity: 88.89%      Conservative: 2
Best Local Similarity: 77.78%    Mismatches: 2
Query Match: 48.45%             Indels: 0
DB: 9                      Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-811-115-2 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTTGGACAGCACCTTC 3012

RESULT 15
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: Intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
```

```
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9
```

```
Alignment Scores:
Pred. No.:      0.0323      Length:      3768
Score:          78.00      Matches:      14
Percent Similarity: 88.89%      Conservative: 2
Best Local Similarity: 77.78%      Mismatches: 2
Query Match:     48.45%      Indels:    0
DB:              10          Gaps:      0
```

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-854-356-9 (1-3768)

```
Qy      13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db      2959 GTGGTCATCCAGAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACACACCTTC 3012
```

Search completed: October 17, 2003, 03:53:07
Job time : 35.0111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:55:19 ; Search time 7.60488 Seconds
(without alignments)
1741.185 Million cell updates/sec

Title: JUNC_SEQ8_SEQ2RES991_

Perfect score: 161

Sequence: 1 ERGCPAQRPASPVTFQNEGLGPPSPMDSTF 30

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/HOLLERAN480/runat_15102003_131913_20570/app_query.fasta_1.4685
-DB=Issued_Patents_NA -QFWT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=pct -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480.ecgn_1.1.403.@runat_15102003_131913_20570 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	84	52.2	3955	1	US-08-229-515A-14
2	84	52.2	3955	1	US-08-645-865-14
3	78	48.4	3768	2	US-08-625-101-1
4	78	48.4	3768	2	US-08-356-786-1
5	78	48.4	4473	2	US-09-048-804-1
6	78	48.4	4473	3	US-09-056-105-26
7	78	48.4	4530	1	US-08-229-515A-9
8	78	48.4	4530	1	US-08-645-865-9
9	78	48.4	4530	4	US-09-167-322-4
10	78	48.4	4530	4	US-09-527-487-1
11	78	48.4	4530	4	US-09-877-177A-11
12	68	42.2	2385	2	US-09-146-283-3

13	68	42.2	2385	3	US-08-579-823A-3	Sequence 3, Appli
14	68	42.2	2385	3	US-09-344-195-3	Sequence 3, Appli
15	66	41.0	153	3	US-08-776-251-3	Sequence 3, Appli
16	66	41.0	201	4	US-09-200-355-3	Sequence 3, Appli
17	66	41.0	816	3	US-08-776-251-10	Sequence 10, Appli
18	59	36.6	1938	4	US-09-547-435-25	Sequence 25, Appli
19	59	36.6	2316	4	US-09-547-435-27	Sequence 27, Appli
20	59	36.6	2604	4	US-09-547-435-23	Sequence 23, Appli
21	59	36.6	3384	4	US-09-547-435-29	Sequence 29, Appli
22	57.5	35.7	2691	4	US-09-020-743-1	Sequence 1, Appli
23	55	34.2	4092	3	US-09-306-595C-5	Sequence 5, Appli
24	55	34.2	4092	4	US-09-925-388-5	Sequence 5, Appli
25	54	33.5	960	3	US-08-651-136C-1	Sequence 1, Appli
26	54	33.5	960	4	US-09-229-911A-1	Sequence 1, Appli
27	54	33.5	5046	4	US-09-548-938A-5	Sequence 5, Appli
28	54	33.5	111282	4	US-09-754-250-3	Sequence 3, Appli
29	53.5	33.2	864	3	US-09-022-983-3	Sequence 3, Appli
30	53.5	33.2	2340	3	US-09-022-983-4	Sequence 4, Appli
31	53.5	33.2	2477	3	US-09-490-692-3	Sequence 3, Appli
32	53.5	33.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
33	53.5	33.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
34	53	32.9	669	4	US-09-252-991A-9807	Sequence 9807, Ap
35	53	32.9	1020	2	US-08-282-137C-1	Sequence 1, Appli
36	53	32.9	1353	4	US-09-252-991A-9985	Sequence 9985, Ap
37	53	32.9	42571	4	US-09-810-347-3	Sequence 3, Appli
38	53	32.9	536165	4	US-09-214-808-1	Sequence 1, Appli
39	52.5	32.6	444	4	US-09-252-991A-15154	Sequence 15154, A
40	52.5	32.6	465	4	US-09-252-991A-14734	Sequence 14734, A
41	52.5	32.6	621	4	US-09-252-991A-15020	Sequence 15020, A
42	52.5	32.6	1107	4	US-09-252-991A-932	Sequence 932, App
43	52.5	32.6	1266	4	US-09-252-991A-14898	Sequence 14898, A
44	52.5	32.6	1527	4	US-09-252-991A-1035	Sequence 1035, Ap
45	52.5	32.6	1531	2	US-08-983-352-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZILUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A-14

Alignment Scores:
Pred. No.: 0.00361 Length: 3955
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-229-515A-14 (1-3955)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
Db 2990 GTGGTCATCCAGAACGAGGACTTGGGCCCATCCAGCCCCATGACAGTACCTTC 3043

RESULT 2
US-08-645-865-14
; Sequence 14, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; TITLE OF INVENTION: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-645-865-14

Alignment Scores:
Pred. No.: 0.00361 Length: 3955
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-645-865-14 (1-3955)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
Db 2990 GTGGTCATCCAGAACGAGGACTTGGGCCCATCCAGCCCCATGACAGTACCTTC 3043
```

```
RESULT 3
US-08-625-101-1
; Sequence 1, Application US/08625101
; Patent No. 5869445
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 0.0285 Length: 3768
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-625-101-1 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCATCCAGTACCTTC 3012

RESULT 4
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
```

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3768
; OTHER INFORMATION: /note= "product = "cerB-b2""
;
; US-08-356-786-1
;
; Alignment Scores:
; Pred. No.: 0.0285 Length: 3768
; Score: 78.00 Matches: 14
; Percent Similarity: 88.89% Conservative: 2
; Best Local Similarity: 77.78% Mismatches: 2
; Query Match: 48.45% Indels: 0
; DB: 2 Gaps: 0
;
; JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-356-786-1 (1-3768)
;
; Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
; ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
; Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGCTCCCTTGGACAGACCTTC 3012
;
; RESULT 5
; US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; HUMAN HER-2 EXPRESSION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748rls LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048,804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: No
;
; US-09-048-804-1
;
; Alignment Scores:
; Pred. No.: 0.0355 Length: 4473
; Score: 78.00 Matches: 14
; Percent Similarity: 88.89% Conservative: 2
; Best Local Similarity: 77.78% Mismatches: 2
; Query Match: 48.45% Indels: 0
; DB: 2 Gaps: 0
;
; JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-048-804-1 (1-4473)
;
; Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
; ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
; Db 3133 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGCTCCCTTGGACAGACCTTC 3186
;
; RESULT 6
; US-09-056-105-26
; Sequence 26, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
; APPLICANT: KIPPS, THOMAS J.
; APPLICANT: WU, YUNQI
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
; TITLE OF INVENTION: PROCESSING
; FILE REFERENCE: 233/221
; CURRENT APPLICATION NUMBER: US/09/056,105
; CURRENT FILING DATE: 1998-04-06
; EARLIER APPLICATION NUMBER: 60/043,467
; EARLIER FILING DATE: 1997-04-10
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
;
; US-09-056-105-26
;
; Alignment Scores:
; Pred. No.: 0.0355 Length: 4473
; Score: 78.00 Matches: 14
; Percent Similarity: 88.89% Conservative: 2
; Best Local Similarity: 77.78% Mismatches: 2
; Query Match: 48.45% Indels: 0
; DB: 3 Gaps: 0
;
; JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-056-105-26 (1-4473)
;
; Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
; ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
; Db 3133 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGCTCCCTTGGACAGACCTTC 3186
;
; RESULT 7
; US-08-229-515A-9
```



```
; Sequence 9, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-229-515A-9

Alignment Scores:
Pred. No.: 0.036 Length: 4530
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-229-515A-9 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCCTTGACAGCACCTTC 3162

RESULT 8
US-08-645-865-9
; Sequence 9, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-645-865-9

Alignment Scores:
Pred. No.: 0.036 Length: 4530
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-645-865-9 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCCTTGACAGCACCTTC 3162

RESULT 9
US-09-167-322-4
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4
Alignment Scores:
Pred. No.: 0.036 Length: 4530
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 4 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-167-322-4 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCCTTC 3162
RESULT 10
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
US-09-527-487-1
Alignment Scores:
Pred. No.: 0.036 Length: 4530
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 4 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-527-487-1 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCCTTC 3162
RESULT 11
US-09-877-177A-11
; Sequence 11, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-New Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11
Alignment Scores:
Pred. No.: 0.036 Length: 4530
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 4 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-877-177A-11 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCCTTC 3162
RESULT 12
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/146,283
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3
Alignment Scores:
Pred. No.: 0.556 Length: 2385
Score: 68.00 Matches: 15
Percent Similarity: 68.97% Conservative: 5
Best Local Similarity: 51.72% Mismatches: 7
Query Match: 42.24% Indels: 2
DB: 2 Gaps: 1
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-146-283-3 (1-2385)
```

```

RESULT 14
US-09-344-195-3
; Sequence 3, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
;           Ruegg, Curtis L.
;           Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-344-195-3

Alignment Scores:
Pred. No.:      0.556      Length:      2385
Score:          68.00     Matches:      15
Percent Similarity: 68.97%   Conservative: 5
Best Local Similarity: 51.72% Mismatches:    7
Query Match:      42.24% Indels:        2
DB:                3       Gaps:         1

JUNC_SEQ08_SEQ02RES991_ (1-30) x US-09-344-195-3 (1-2385)
Qy      1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnG
      : ::::: ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db      1925 GACAAGGGCTGCCCGCCGACGAGACGACGACGCCTCTGACG-----TCCC
Qy      21 GlyProSerSerPrometAspSerThr 29
      : ||||| |||||
Db      1979 CCCGCCCGCTCGCCCCACGCCCCAGCACCA 2005

RESULT 15
US-08-776-251-3
; Sequence 3, Application US/08776251

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 28.4941 Seconds
(without alignments)
2842.104 Million cell updates/sec

Title: JUNC_SEQ8_SEQ2RES991_

Perfect score: 161

Sequence: 1 ERGCPAEQASRPVTFQNEGLPSPMDSTF 30

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cqn2.1/USPFO_spool_p/HOLLERAN480/runat_15102003_131912_20526/app_query.fasta_1.4685
-DB=N_Geneseq_19Jun03 -QMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480.ecgn_1.1.1758.#runat_15102003_131912_20526 -NCPU=6 -ICPU=3
-NO_MAP -LARGUEURY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	153	95.0	2763	24	ABA92252 Mouse Her-2/neu ex
2	153	95.0	2781	24	ABA92253 Her-2/neu extracel
3	84	52.2	3771	21	AAA9737 Mouse Her-2/neu CD
4	84	52.2	3771	22	AAH42210 Nucleotide sequenc
5	84	52.2	3771	24	ABA92251 Mouse Her-2/neu CD
6	84	52.2	3955	16	AAT01590 Rat neu promoter.
7	84	52.2	3955	21	AAA9753 Rat HER-2/neu prot
8	78	48.4	1115	21	AAF21778 Human breast and o
9	78	48.4	1755	24	AAD32746 Human cDNA for the
10	78	48.4	1767	24	AAD32747 Human cDNA for the
11	78	48.4	1773	24	AAD32747 Human cDNA for the
12	78	48.4	1806	24	AAD32745 Human cDNA for the
13	78	48.4	3600	21	AAA89736 Human HER-2/neu co
14	78	48.4	3678	24	ABK86207 cDNA encoding huma
15	78	48.4	3768	17	AAT40739 HER-2/neu oncogene
16	78	48.4	3768	20	AAK01912 Human HER-2/neu on
17	78	48.4	3768	21	AAA09455 Human heregulin 2
18	78	48.4	3768	22	AAH23392 Human HER-2/neu pr
19	78	48.4	3768	24	AB235744 Human ERBB2 polynu
20	78	48.4	3768	24	ABX09987 Human ERBB2 DNA fr
21	78	48.4	3768	24	AAD43935 Human HER-2 cDNA.
22	78	48.4	3768	24	AAD43986 Human Her2 antigen
23	78	48.4	3768	24	ABV78168 Human ERBB2 DNA SE
24	78	48.4	3768	24	AAD32743 Human Her-2/neu pr
25	78	48.4	3768	24	ABA92250 Human Her-2/neu CD
26	78	48.4	3768	24	ABK10730 Human Her-2/neu DN
27	78	48.4	3768	24	ABL1709 Human polynucleoti
28	78	48.4	3768	24	ABK14058 Human HER2 (ERBB2)
29	78	48.4	4299	14	AAQ46083 Sequence encoding the
30	78	48.4	4472	21	AAAI4812 cDNA encoding the
31	78	48.4	4473	19	ABO76220 Human tumour antiq
32	78	48.4	4473	20	AZ31071 HER-2 nucleic acid
33	78	48.4	4473	24	ABZ34969 Human gene express
34	78	48.4	4473	24	AAD38904 Human Her-2 DNA.
35	78	48.4	4530	16	AAT01585 Her-2/neu (ERBB2/c
36	78	48.4	4530	18	AAT71253 Human HER2 gene.
37	78	48.4	4530	21	AAZ60815 Nucleotide sequenc
38	78	48.4	4530	22	AAD19731 Human tyrosine kin
39	78	48.4	4530	24	ABZ35012 Human gene express
40	78	48.4	4530	24	ABV94128 Breast carcinoma r
41	78	48.4	4530	24	ABN85585 Human HER2-neu SEQ
42	78	48.4	4530	24	ABK83918 Human cDNA differe
43	78	48.4	4530	25	ACC50139 Breast cancer asso
44	78	48.4	4530	25	ABQ83856 Human Her2/Neu enc
45	78	48.4	9274	22	AAF24297 HER2 transgene pla

ALIGNMENTS

RESULT 1
ABA92252
ID ABA92252 standard; cDNA: 2763 BP.
XX
XX ABA92252;
AC
XX
XX 17-JUN-2002 (first entry)
DT
XX
DE Mouse Her-2/neu extracellular-phosphorylation domain fusion cDNA.
XX
KW Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
KW receptor; mouse; gene therapy; gene; ss.
XX
XX Mus musculus.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..2763

```

FT      /*tag= a
XX      /product= "Her-2/neu ECD-PD fusion"
XX
XX      WO200212341-A2.
XX
XX      14-FEB-2002.
XX
XX      03-AUG-2001; 2001WO-US24283.
XX
XX      03-AUG-2000; 2000US-0632507.
XX
XX      (CORI-) CORIXA CORP.
XX      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Cheever MA, Gheysen D;
XX
XX      WPI; 2002-241743/29.
XX      P-PSDB; AAMS1152.
XX
XX      Her-2/neu fusion protein for treating or preventing cancer by eliciting
XX      or enhancing an immune response to the protein, has Her-2/neu
XX      extracellular domain fused to Her-2/neu intracellular or
XX      phosphorylation domain
XX
XX      Disclosure; Fig 23; 141pp; English.
XX
XX      The present sequence is that of cDNA encoding a fusion between
XX      the extracellular domain (ECD) and the phosphorylation domain (PD)
XX      murine Her-2/neu oncoprotein (see AAMS1151). The Her-2/neu gene is
XX      amplified and the oncoprotein is overexpressed in a variety of
XX      human cancers, including breast, ovarian, colon, lung and prostate
XX      cancer. Her-2/neu overexpression correlates with a poor prognosis
XX      in breast and ovarian cancers. The invention provides Her-2/neu
XX      fusion proteins, nucleic acids encoding them, viral vectors, and
XX      vaccines comprising the fusion proteins or nucleic acid molecules.
XX      In preferred fusion proteins, the ECD of Her-2/neu is fused to a
XX      Her-2/neu intracellular domain or PD (or its DeltapB fragment). An
XX      immune response to Her-2/neu protein is elicited or enhanced by
XX      administering the fusion protein in the form of a vaccine, or by
XX      transfecting cells of an animal ex vivo with a nucleic acid
XX      encoding the fusion protein, and delivering the transfected cells
XX      to the animal. The fusion proteins, nucleic acids, and isolated
XX      cancer, especially breast, ovarian, colon, lung or prostate cancer
XX      in a patient. T cells that specifically react with a Her-2/neu
XX      fusion protein can be used to remove tumour cells from a sample in
XX      order to inhibit the development of cancer in a patient.
XX
XX      Sequence 2763 BP; 571 A; 855 C; 772 G; 565 T; 0 other;

Alignment Scores:
Pred. No.:      1.58e-12      Length:      2763
Score:          153.00      Matches:      29
Percent Similarity: 96.67%      Conservative: 0
Best Local Similarity: 96.67%      Mismatches: 1
Query Match:    95.03%      Indels:      0
DB:             24      Gaps:        0

JUNC_SEQ8_SEQ2RES991_ (1-30) x ABA92252 (1-2763)

Qy      1  GLuArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Dy      1918 GAACGAGGCTGCCAGCAGACGAGAGCCAGCCAGTGCAGTCTCAGAACGAGGACTTA 1977
Qy      21  GlyProSerSerProMetAspSerThrPhe 30
Dy      1978 GGCCCCCTCCAGCCCCATGGACAGCACCCTC 2007

RESULT 2
ID      ABA92253 standard; cDNA; 2781 BP.
XX
XX      ABA92253;
XX
XX      ABA92253;
XX
XX      17-JUN-2002 (first entry)
XX
XX      Her-2/neu extracellular-phosphorylation domain-Tcp0 fusion cDNA.
XX
XX      Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
XX      receptor; Tcp0; mouse; gene therapy; gene; ss.
XX
XX      Chimeric - Mus musculus.
XX      Chimeric - Unidentified.
XX
XX      Key      Location/Qualifiers
XX      CDS      1..2781
XX      /*tag= a
XX      /product= "Her-2/neu ECD-PD-Tcp0 fusion"
XX
XX      WO200212341-A2.
XX
XX      14-FEB-2002.
XX
XX      03-AUG-2001; 2001WO-US24283.
XX
XX      03-AUG-2000; 2000US-0632507.
XX
XX      (CORI-) CORIXA CORP.
XX      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Cheever MA, Gheysen D;
XX
XX      WPI; 2002-241743/29.
XX      P-PSDB; AAMS1153.
XX
XX      Her-2/neu fusion protein for treating or preventing cancer by eliciting
XX      or enhancing an immune response to the protein, has Her-2/neu
XX      extracellular domain fused to Her-2/neu intracellular or
XX      phosphorylation domain
XX
XX      Disclosure; Fig 25; 141pp; English.
XX
XX      The present sequence is that of cDNA encoding a fusion between
XX      the extracellular domain (ECD) and the phosphorylation domain (PD)
XX      murine Her-2/neu oncoprotein (see AAMS1151) plus a C-terminal
XX      Tcp0 motif that improves immunogenicity. The Her-2/neu gene is
XX      amplified and the oncoprotein is overexpressed in a variety of
XX      human cancers, including breast, ovarian, colon, lung and prostate
XX      cancer. Her-2/neu overexpression correlates with a poor prognosis
XX      in breast and ovarian cancers. The invention provides Her-2/neu
XX      fusion proteins, nucleic acids encoding them, viral vectors, and
XX      vaccines comprising the fusion proteins or nucleic acid molecules.
XX      In preferred fusion proteins, the ECD of Her-2/neu is fused to a
XX      Her-2/neu intracellular domain or PD (or its DeltapB fragment). An
XX      immune response to Her-2/neu protein is elicited or enhanced by
XX      administering the fusion protein in the form of a vaccine, or by
XX      transfecting cells of an animal ex vivo with a nucleic acid
XX      encoding the fusion protein, and delivering the transfected cells
XX      to the animal. The fusion proteins, nucleic acids, and isolated
XX      cancer, especially breast, ovarian, colon, lung or prostate cancer
XX      in a patient. T cells that specifically react with a Her-2/neu
XX      fusion protein can be used to remove tumour cells from a sample in
XX      order to inhibit the development of cancer in a patient.
XX
XX      Sequence 2781 BP; 574 A; 859 C; 779 G; 569 T; 0 other;

Alignment Scores:
Pred. No.:      1.59e-12      Length:      2781
Score:          153.00      Matches:      29
Percent Similarity: 96.67%      Conservative: 0
Best Local Similarity: 96.67%      Mismatches: 1
Query Match:    95.03%      Indels:      0
DB:             24      Gaps:        0

JUNC_SEQ8_SEQ2RES991_ (1-30) x ABA92253 (1-2781)

```

```
QY 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACAGAGAGCCAGCCAGTCAGTCTCAGAACGAGGACTTA 1977
QY 21 GlyProSerSerProMetAspSerThrPhe 30
Db 1978 GGCCCTCCAGCCCATGGACAGCACCTTC 2007

RESULT 3
AAA89737
ID AAA89737 standard; cDNA; 3771 BP.
XX AC AAA89737;
XX DT 12-JAN-2001 (first entry)
XX DE Mouse Her-2/neu cDNA.
XX KW Mouse; Her-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
XX KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
XX KW colon cancer; ss.
XX OS Mus sp.
XX PN WO200044899-A1.
XX PD 03-AUG-2000.
XX PF 28-JAN-2000; 2000WO-US02164.
XX PR 29-JAN-1999; 99US-0117976.
XX PA (CORI-) CORIXA CORP.
XX PI (SMIK ) SMITHKLINE BEECHAM.
XX PT Cheever MA, Gheysen D;
XX DR WPI; 2000-505976/45.
XX DR P-PSDB; AAB21206.
XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins
XX PT useful for vaccinating against breast, ovarian, colon, lung and
XX PT prostate cancers -
XX PS Disclosure; Fig 19; 128pp; English.
XX CC The present sequence encodes the mouse Her-2/neu protein. Her-2/neu is a
XX CC member of the tyrosine kinase family of receptor-like glycoproteins and
XX CC shows homology to the epidermal growth factor receptor (EGFR). It
XX CC probably plays a part in cell growth and/or differentiation. The
XX CC Her-2/neu gene is an oncogene. An Her-2/neu fusion protein comprising a
XX CC Her-2/neu extracellular domain fused to a Her-2/neu phosphorylation
XX CC domain may be used to treat or prevent cancer by eliciting or
XX CC enhancing an immune response to the Her-2/neu protein. It may be used
XX CC to treat malignancies such as breast, ovarian, colon, lung and
XX CC prostate cancers, and may be used as an antigen to vaccinate against
XX CC these neoplasias.
XX SQ Sequence 3771 BP; 808 A; 1105 C; 1068 G; 790 T; 0 other;

Alignment Scores:
Pred. No.: 0.0342 Length: 3771
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AAA89737 (1-3771)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 1918 GAACGAGGCTGCCAGCAGACAGAGAGCCAGTCAGTCTCAGAACGAGGACTTA 1977
```

```
Db 2962 GTGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCATGGACAGCACCTTC 3015

RESULT 4
AAH42210
ID AAH42210 standard; cDNA; 3771 BP.
XX AC AAH42210;
XX DT 17-SEP-2001 (first entry)
XX DE Nucleotide sequence of a murine neu polypeptide from C57B16 mice.
XX KW Neu polypeptide; Her-2; cancer; breast cancer; T cell expansion;
XX KW vaccine; ss.
XX OS Mus sp.
XX FH Key Location/Qualifiers
XX FT 1..3771
XX FT /*tag= a
XX FT /product= "neu polypeptide"
XX PN WO200148205-A2.
XX PD 05-JUL-2001.
XX PF 29-DEC-2000; 2000WO-US35648.
XX PR 29-DEC-1999; 99US-0474382.
XX PA (CORI-) CORIXA CORP.
XX PI Spies AG;
XX DR WPI; 2001-441718/47.
XX DR P-PSDB; AAG62860.
XX PT Novel isolated murine homolog of human Her-2/neu useful for inhibiting
XX PT development of cancer, preferably breast cancer in a patient -
XX PS Claim 4; Fig 1A-B; 53pp; English.
XX CC The present sequence encodes a murine neu polypeptide. It is a murine
XX CC homologue of human Her-2/neu. Neu polypeptides and polynucleotides are
XX CC useful for inhibiting the development of a cancer, preferably breast
XX CC cancer in a patient. They are also useful for stimulating and/or
XX CC expanding T cells specific for Her-2/neu. They are useful for detecting
XX CC and treating Her-2/neu malignancies. Neu polypeptides are useful
XX CC for producing vaccines.
XX SQ Sequence 3771 BP; 808 A; 1105 C; 1068 G; 790 T; 0 other;

Alignment Scores:
Pred. No.: 0.0342 Length: 3771
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 22 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AAH42210 (1-3771)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 2962 GTGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCATGGACAGCACCTTC 3015

RESULT 5
ABA92251
ID ABA92251 standard; cDNA; 3771 BP.
XX AC ABA92251;
XX DT 17-JUN-2002 (first entry)
```


XX Mouse Her-2/neu cDNA.
 DE Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
 KW receptor; mouse; gene therapy; gene; ss.
 XX Mus musculus.
 OS
 XX Key Location/Qualifiers
 FT CDS 1..3771
 FT /*tag= a
 FT /product= "Her-2/neu"
 XX WO200212341-A2.
 XX
 XX 14-FEB-2002.
 XX
 XX 03-AUG-2001; 2001WO-US24283.
 XX
 XX 03-AUG-2000; 2000US-0632507.
 XX
 XX (CORI-) CORIXA CORP.
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 XX Cheever MA, Gheysen D;
 PI WPI; 2002-241743/29.
 DR P-PSDB; AAM51151.
 DR
 XX
 XX Her-2/neu fusion protein for treating or preventing cancer by eliciting
 PT or enhancing an immune response to the protein, has Her-2/neu
 PT extracellular domain fused to Her-2/neu intracellular or
 PT phosphorylation domain
 XX
 XX Disclosure; Fig 19; 141pp; English.
 PS
 XX The present sequence is that of murine Her-2/neu oncogene cDNA.
 CC The cDNA can be obtained by PCR amplification (see ABA92255-56)
 CC from mouse brain RNA. The Her-2/neu oncogene encodes a p185
 CC glycoprotein (see AAM51151). The Her-2/neu gene is amplified and
 CC p185 is overexpressed in a variety of cancers, including breast,
 CC ovarian, colon, lung and prostate cancer. Her-2/neu is a member
 CC of the tyrosine kinase family of receptor-like glycoproteins.
 CC Its overexpression correlates with a poor prognosis in breast and
 CC ovarian cancers. The invention provides Her-2/neu fusion
 CC proteins, nucleic acids encoding them, viral vectors, and vaccines
 CC comprising the fusion proteins or nucleic acid molecules. In
 CC preferred fusion proteins, the extracellular domain of a Her-2/neu
 CC protein is fused to a Her-2/neu intracellular domain or
 CC phosphorylation domain (or its deltaTPD fragment). An immune
 CC response to Her-2/neu protein is elicited or enhanced by
 CC administering the fusion protein in the form of a vaccine, or by
 CC transfecting cells of an animal ex vivo with a nucleic acid
 CC encoding the fusion protein, and delivering the transfected cells
 CC to the animal. The fusion proteins, nucleic acids, and isolated
 CC specific T-cells are useful for inhibiting the development of a
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer
 CC in a patient. T cells that specifically react with a Her-2/neu
 CC fusion protein can be used to remove tumour cells from a sample in
 CC order to inhibit the development of cancer in a patient.
 XX
 SQ Sequence 3771 BP; 808 A; 1105 C; 1068 G; 790 T; 0 other;

Alignment Scores:
 Pred. No.: 0.0342 Length: 3771
 Score: 84.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 52.17% Indels: 0
 DB: 24 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x ABA92251 (1-3771)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
 Db |||||
 2962 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCCATGACAGCACCTTC 3015
 RESULT 6
 AAT01590
 ID AAT01590 standard; DNA; 3955 BP.
 XX
 AC AAT01590;
 XX
 DT 21-APR-1996 (first entry)
 XX
 DE Rat neu promoter.
 XX
 KW Her-2/neu; c-erbB-2; oncogene; Erb-B2 promoter; DNA binding protein;
 KW HPBF; Erb-B2 promoter binding protein; tumour enhancer factor;
 KW breast cancer diagnosis; prognosis; antisense oligonucleotide;
 KW retro virus vector; gene therapy vector; ss.
 XX
 OS Rattus rattus.
 XX
 PN WO9528485-A1.
 XX
 PD 26-OCT-1995.
 XX
 XX 19-APR-1995; 95WO-US04953.
 XX
 PR 19-APR-1994; 94US-0229515.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Raziuddin F, Sarkar FH;
 PI
 XX WPI; 1995-373800/48.
 DR
 XX

New purified protein binding to the ERBB2 gene promoter - to induce cell proliferation, diagnostic of breast cancer, also related antibodies, nucleic acid, assays and methods for screening inhibitors.

Disclosure; Page 57-59; 69pp; English.

The sequence represents one of the promoters of the rat neu gene; a second rat neu gene promoter is in AAT01589. These promoter are compared with the human Erb-B2 promoters (AAT01587-88) and the mouse neu promoter (AAT01591). The human Erb-B2 gene is one of the primary genes responsible for the transition of normal breast epithelial cells towards carcinoma in situ and the subsequent development of invasive and metastatic cancer. The Erb-B2 promoter-binding protein, HPBF (see AAR77093-94), induces cell division on binding to the promoter. In a method for greater success in early identification and treatment of breast cancer, the initiation step for Erb-B2 gene activity is identified. This method involves determining the presence of HPBF in a biopsy from the subject, where the presence of HPBF (relative to its absence in a normal control) indicates the presence of cancer and a decreased chance of long-term survival. Binding of HPBF to the promoter can be inhibited using antisense oligonucleotides or a non-genomic nucleic acid that binds to HPBF; these oligos can be expressed from retro virus or other gene therapy vectors.

SQ Sequence 3955 BP; 842 A; 1147 C; 1136 G; 830 T; 0 other;

Alignment Scores:
 Pred. No.: 0.0362 Length: 3955
 Score: 84.00 Matches: 16
 Percent Similarity: 88.89% Conservative: 0
 Best Local Similarity: 88.89% Mismatches: 2
 Query Match: 52.17% Indels: 0
 DB: 16 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AAT01590 (1-3955)

JUNC_SEQ8_SEQ2RES991_ (1-30) x AAA89753 (1-3955)	
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSer:SerProMetAspSerThrPhe 30	
Db 2990 GTGGTCATCCAGACGAGGACTTGGGCCCATCCAGCCCATGGACAGTACCTTC 3043	
RESULT 8	
AAAF21778	
ID AAF21778 standard; DNA; 1115 BP.	
XX AAF21778;	
XX 27-MAR-2001 (first entry)	
XX Human breast and ovarian cancer associated antigen gene SEQ ID 165.	
XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;	
KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;	
KW antidiabetic; antiinflammatory; antitumor; antiparasitic; anticonvulsant;	
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;	
KW Addison's disease; allergy; autoimmune haemolytic anaemia;	
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;	
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;	
KW cardiovascular disorder; wound healing; neurological disease; ds.	
XX Homo sapiens.	
OS WO200055173-A1.	
XX 21-SEP-2000.	
XX 08-MAR-2000; 2000WO-US05881.	
XX 12-MAR-1999; 99US-0124270.	
XX (HUMA-) HUMAN GENOME SCI INC.	
XX Rosen CA, Ruben SM;	
XX WPI; 2000-611515/58.	
XX P-PSDB; AAF58875.	
XX New human breast and ovarian cancer associated gene sequences and th	
PT polypeptides encoded by these genes, useful in the prevention,	
PT treatment and diagnosis of cancer, immune disorders, cardiovascular	
PT disorders and neurological diseases -	
XX Claim 1; Page 604; 1299pp; English.	
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human	
CC proteins AAF58711 - AAF59128. The DNA and protein sequences are	
CC associated with breast and ovarian cancer. Included in the invention	
CC sequences AAF22032 - AAF22040 and AAF59129 which are used in the	
CC isolation and characterisation of the DNA and protein sequences of t	
CC invention. The breast and ovarian cancer associated DNA, protein, ag	
CC or antagonist sequences exhibit cytostatic; immunosuppressive;	
CC neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;	
CC antibacterial; antiinflammatory; antitumor; vulnary; anticonvulsant;	
CC antibacterial; antifungal; antiparasitic and cardiant activity. The	
CC polynucleotide and protein sequences are used in the diagnosis of ca	
CC particularly breast and ovarian cancer. The nucleic acid sequences,	
CC proteins, agonists and antagonists may also be used in the diagnosis,	
CC prevention and treatment of immune disorders e.g. Addison's disease,	
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,	
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid	
CC arthritis and ulcerative colitis; cardiovascular disorders such as	
CC myocardial ischaemias; wound healing; neurological diseases such as	
CC cerebral anoxia and epilepsy; and infectious diseases.	
XX Sequence 1115 BP; 210 A; 336 G; 222 T; 9 other;	
SQ Alignment Scores:	
Pred. No.: 0.0601 Length: 1115	

Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AAF21778 (1-1115)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrpHe 30
||| |||||||||||||||||||:::|||||:::|||||||||
Db 446 GTGGTCATCCAGATGAGCAGCTTGGCCCGACGAGTCCCTTGACACACCTTC 499

RESULT 9
AAD32746
ID AAD32746 standard; cDNA; 1755 BP.
XX
AC AAD32746;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human cDNA for the clone HICD_native_coding_region.
XX
KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
XX human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
OS Homo sapiens.

Key Location/Qualifiers
FH 1..1755
FT CDS
FT FT
FT /tag= a
FT /product= "Human protein encoded by cDNA for the clone
HICD_native_coding_region"
FT /trans_except= (pos:1741..1752, aa:Leu-Glu)
FT /note= "CDS does not include stop codon"
FT /partial
XX
PN WO200214503-A2.
XX
XX 21-FEB-2002.
XX
PF 14-AUG-2001; 2001WO-US41733.
XX
PR 14-AUG-2000; 2000US-225152P.
XX 28-SEP-2000; 2000US-236428P.
XX 21-FEB-2001; 2001US-270520P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
PI McNeill PD, Vedvick TS;
XX
XX WPI; 2002-280758/32.
DR P-PSDB; AAE20483.
XX
PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,
PT prevention and diagnosis of cancer, preferably breast cancer -
XX
PS Example 5; Page 118-119; 129pp; English.
XX
CC The invention relates to an isolated Her-2/Neu polypeptide composition
CC effective for eliciting an immune response. The invention is useful for
CC eliciting an immune response in a patient, where the patient is human
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
CC The composition is useful for the therapy and diagnosis of cancer,
CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
CC and other compositions for the diagnosis, prevention and treatment of
CC human malignancies, for stimulating and/or expanding T cells specific for
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
CC patient. The invention is useful for stimulating a T cell response in a
CC human patient, as probe or primer for nucleic acid hybridisation, to
CC selectively form duplex molecules with complementary stretches of the
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
CC length gene from a suitable library, and to direct expression of a

polypeptide in appropriate host cells. The composition is useful in
prophylactic or therapeutic applications and for the treatment of cancer
preferably for the immunotherapy of breast cancer and other Her-2/Neu-
associated malignancies. The invention is useful in gene therapy. The
present sequence is human cDNA for the clone HICD_native_coding_region.

Sequence 1755 BP; 376 A; 517 C; 531 G; 331 T; 0 other;

Alignment Scores:
Pred. No.: 0.104 Length: 1755
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AAD32746 (1-1755)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrpHe 30
||| |||||||||||||||||||:::|||||:::|||||||||
Db 937 GTGGTCATCCAGATGAGCAGCTTGGCCCGACGAGTCCCTTGACACACCTTC 990

RESULT 10
AAD32744
ID AAD32744 standard; cDNA; 1767 BP.
XX
AC AAD32744;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human cDNA for the clone HICD_CT_His_coding_region.
XX
KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
XX human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
OS Homo sapiens.

Key Location/Qualifiers
FH 1..1764
FT CDS
FT /*tag= a
FT /product= "Human protein encoded by cDNA for the clone
HICD_CT_His_coding_region"
FT
FT
XX PN WO200214503-A2.
XX PD 21-FEB-2002.
XX
XX PF 14-AUG-2001; 2001WO-US41733.
XX
PR 14-AUG-2000; 2000US-225152P.
XX 28-SEP-2000; 2000US-236428P.
XX 21-FEB-2001; 2001US-270520P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
PI McNeill PD, Vedvick TS;
XX
XX WPI; 2002-280758/32.
DR P-PSDB; AAE20481.
XX
PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,
PT prevention and diagnosis of cancer, preferably breast cancer -
XX
PS Example 5; Page 117-118; 129pp; English.
XX
CC The invention relates to an isolated Her-2/Neu polypeptide composition
CC effective for eliciting an immune response. The invention is useful for
CC eliciting an immune response in a patient, where the patient is human
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
CC The composition is useful for the therapy and diagnosis of cancer,
CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
CC and other compositions for the diagnosis, prevention and treatment of
CC human malignancies, for stimulating and/or expanding T cells specific for
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
CC patient. The invention is useful for stimulating a T cell response in a
CC human patient, as probe or primer for nucleic acid hybridisation, to
CC selectively form duplex molecules with complementary stretches of the
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
CC length gene from a suitable library, and to direct expression of a

CC human malignancies, for stimulating and/or expanding T cells specific for
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
 CC patient. The invention is useful for stimulating a T cell response in a
 CC human patient, as probe or primer for nucleic acid hybridisation, to
 CC selectively form duplex molecules with complementary stretches of the
 CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
 CC length gene from a suitable library, and to direct expression of a
 CC polypeptide in appropriate host cells. The composition is useful in
 CC prophylactic or therapeutic applications and for the treatment of cancer,
 CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
 CC associated malignancies. The invention is useful in gene therapy. The
 CC present sequence is human cDNA for the clone HICD_CT_His_coding_region.
 XX
 SQ Sequence 1767 BP; 381 A; 521 C; 529 G; 336 T; 0 other;

Alignment Scores:
 Pred. No.: 0.105 Length: 1767
 Score: 78.00 Matches: 14
 Percent Similarity: 88.89% Conservative: 2
 Best Local Similarity: 77.78% Mismatches: 2
 Query Match: 48.45% Indels: 0
 DB: 24 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AAD32744 (1-1767)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
 Db 937 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 990

RESULT 11
 AAD32747

ID AAD32747 standard; cDNA; 1773 BP.

AC AAD32747;

XX 01-JUL-2002 (first entry)

XX Human cDNA for the clone HICD_in_ppDM_coding_sequence.

XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 XX human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1770

XX /*tag= a

XX /product= "Human protein encoded by cDNA for the clone
 HICD_in_ppDM_coding_sequence"

XX WO200214503-A2.

XX 21-FEB-2002.

XX 14-AUG-2001; 2001WO-US41733.

XX 14-AUG-2000; 2000US-225152P.

XX 28-SEP-2000; 2000US-236428P.

XX 21-FEB-2001; 2001US-270520P.

XX (CORI-) CORIXA CORP.

XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;

XX McNeill PD, Vedvick TS;

XX WPI; 2002-280758/32.

XX P-PSDB; AAE20484.

XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,
 XX prevention and diagnosis of cancer, preferably breast cancer -

XX Example 5; Page 119; 129pp; English.

CC The invention relates to an isolated Her-2/Neu polypeptide composition
 CC effective for eliciting an immune response. The invention is useful for
 CC eliciting an immune response in a patient, where the patient is human
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
 CC The composition is useful for the therapy and diagnosis of cancer,
 CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
 CC and other compositions for the diagnosis, prevention and treatment of
 CC human malignancies, for stimulating and/or expanding T cells specific for
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
 CC patient. The invention is useful for stimulating a T cell response in a
 CC human patient, as probe or primer for nucleic acid hybridisation, to
 CC selectively form duplex molecules with complementary stretches of the
 CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
 CC length gene from a suitable library, and to direct expression of a
 CC polypeptide in appropriate host cells. The composition is useful in
 CC prophylactic or therapeutic applications and for the treatment of cancer,
 CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
 CC associated malignancies. The invention is useful in gene therapy. The
 CC present sequence is human cDNA for the clone HICD_in_ppDM_coding_
 XX sequence.
 XX SQ Sequence 1773 BP; 383 A; 528 C; 530 G; 332 T; 0 other;

Alignment Scores:

Pred. No.: 0.105 Length: 1773
 Score: 78.00 Matches: 14
 Percent Similarity: 88.89% Conservative: 2
 Best Local Similarity: 77.78% Mismatches: 2
 Query Match: 48.45% Indels: 0
 DB: 24 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AAD32747 (1-1773)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
 Db 961 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 1014

RESULT 12
 AAD32745

ID AAD32745 standard; cDNA; 1806 BP.

XX AAD32745;

XX 01-JUL-2002 (first entry)

XX Human cDNA for the clone HICD_plus_8_HIS.

XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 XX human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1803

XX /*tag= a

XX /product= "Human protein encoded by cDNA for the clone
 HICD_plus_8_HIS"

XX /transl_except= (pos:1543..1545, aa:Pro)

XX WO200214503-A2.

XX 21-FEB-2002.

XX 14-AUG-2001; 2001WO-US41733.

XX 14-AUG-2000; 2000US-225152P.

XX 28-SEP-2000; 2000US-236428P.

XX 21-FEB-2001; 2001US-270520P.

XX (CORI-) CORIXA CORP.

XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;

XX McNeill PD, Vedvick TS;

```
XX WPI; 2002-280758/32.
DR P-PSDB; AAE20482.
XX
XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,
PT prevention and diagnosis of cancer, preferably breast cancer -
XX
XX Example 5; Page 118; 129pp; English.
XX
XX The invention relates to an isolated Her-2/Neu polypeptide composition
CC effective for eliciting an immune response. The invention is useful for
CC eliciting an immune response in a patient, where the patient is human
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
CC The composition is useful for the therapy and diagnosis of cancer,
CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
CC and other compositions for the diagnosis, prevention and treatment of
CC human malignancies, for stimulating and/or expanding T cells specific for
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
CC patient. The invention is useful for stimulating a T cell response in a
CC human patient, as probe or primer for nucleic acid hybridisation, to
CC selectively form duplex molecules with complementary stretches of the
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
CC length gene from a suitable library, and to direct expression of a
CC polypeptide in appropriate host cells. The composition is useful in
CC prophylactic or therapeutic applications and for the treatment of cancer,
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
CC associated malignancies. The invention is useful in gene therapy. The
CC present sequence is human cDNA for the clone HICD_plus_8_HIS.
XX
XX Sequence 1806 BP; 391 A; 530 C; 544 G; 341 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 0.108 Length: 1806
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 0
Query Match: 48.45% Indels: 0
DB: 24 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x AAD32745 (1-1806)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
Db 994 GTGGTCATCCAGAATGAGGACTTGGCCCGCCAGTCCTTGGACAGCACCTTC 1047
RESULT 13
AAA89736
ID AAA89736 standard; DNA; 3600 BP.
XX
XX AAA89736;
AC
XX
XX 12-JAN-2001 (first entry)
DT
XX
XX Human HER-2/neu coding sequence.
DE
XX
XX Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
KW colon cancer; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1..3600
FT CDS
FT /*tag= a
FT /product= "HER-2/neu protein"
XX
XX WO200044899-A1.
FN
XX
XX 03-AUG-2000.
PD
XX
XX 28-JAN-2000; 2000WO-US02164.
PF
XX
XX 29-JAN-1999; 99US-0117976.
PR
```

```
XX (CORI-) CORIXA CORP.
PA (SMIK ) SMITHKLINE BEECHAM.
XX
XX Cheever MA, Gheysen D;
PI
XX WPI; 2000-505976/45.
DR
DR P-PSDB; AAB21198, AAB21208.
XX
XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
XX Disclosure; Fig 15; 128pp; English.
XX
XX The present sequence encodes the human HER-2/neu protein. HER-2/neu is
CC a member of the tyrosine kinase family of receptor-like glycoproteins
CC and shows homology to the epidermal growth factor receptor (EGFR). It
CC probably plays a part in cell growth and/or differentiation. The
CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising
CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
XX
XX Sequence 3600 BP; 723 A; 1108 C; 1075 G; 594 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 0.248 Length: 3600
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 0
Query Match: 48.45% Indels: 0
DB: 21 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x AAA89736 (1-3600)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAATGAGGACTTGGCCCGCCAGTCCTTGGACAGCACCTTC 3012
RESULT 14
ABK86207
ID ABK86207 standard; cDNA; 3678 BP.
XX
XX ABK86207;
AC
XX
XX 24-SEP-2002 (first entry)
DT
XX
XX cDNA encoding human breast cancer antigen, Her2 variant.
DE
XX
XX Human; Her2; cytostatic; antiviral; immunostimulant;
KW cell-mediated immune response; tumour; breast cancer;
KW virus infection; prostate cancer; colorectal cancer; pancreatic cancer;
KW lymphoma; leukaemia; hepatitis virus; lentivirus; herpesvirus;
KW human immunodeficiency virus; HIV; flavivirus; pestivirus; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 7..3678
FT CDS
FT /*tag= a
FT /product= "Breast cancer antigen Her2 variant"
XX
XX WO200240059-A2.
FN
XX
XX 23-MAY-2002.
PD
XX
XX 01-NOV-2001; 2001WO-US45626.
PF
XX
XX 01-NOV-2000; 2000US-0704232.
PR
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:09:13 : Search time 432.026 Seconds
(without alignments)
2840.777 Million cell updates/sec

Title: JUNC_SEQ8_SEQ2RES991_
Perfect score: 161
Sequence: 1 ERGCPAEQRAQSPVTFQNEIDLGSPSPMDSTF 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=framer-p2n.model -DEV=xlp
-O/cgn2_1/USPTO_spool_p/HOLLERAM480/runat_15102003_131912_20536/app_query.fasta_1.4685
-DB=GenEmbl -QEXT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=HOLLERAM480 -SCGN 1.1 2724 &runat_15102003_131912_20536
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rnd.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	153	95.0	2763	6	AX380942 Sequence
2	153	95.0	2781	6	AX380944 Sequence
3	84	52.2	2083	10	BC027080 Mus muscu
4	84	52.2	3771	6	AX189662 Sequence
5	84	52.2	3771	6	AX380925 Sequence
6	84	52.2	3955	6	AX380924 Sequence
7	84	52.2	3955	6	I21129 Sequence 14
8	84	52.2	3955	6	I59750 Sequence 14
9	84	52.2	3955	10	RNNEUR X03362 Rat mRNA fo
10	84	52.2	4694	10	BC046811 Mus muscu
11	84	52.2	4695	10	BC053078 Mus muscu
12	84	52.2	4727	10	AY116182 Rattus no
13	81.5	50.6	186134	10	AL591390 Mouse DNA
14	81.5	50.6	214019	2	AC064803 Mus muscu
15	81	50.3	4062	10	HAMNEU D16295 Mesocricatu
16	78	48.4	1755	6	AX384609 Sequence
17	78	48.4	1767	6	AX384607 Sequence
18	78	48.4	1773	6	AX384610 Sequence
19	78	48.4	1806	6	AX384608 Sequence
20	78	48.4	3678	6	AX505114 Sequence
21	78	48.4	3768	6	ARO34479 Sequence
22	78	48.4	3768	6	AX060704 Sequence
23	78	48.4	3768	6	AX201817 Sequence
24	78	48.4	3768	6	AX380923 Sequence
25	78	48.4	3768	6	AX384604 Sequence
26	78	48.4	3768	6	AX465456 Sequence
27	78	48.4	3768	6	AX467229 Sequence
28	78	48.4	3768	6	AX481438 Sequence
29	78	48.4	3780	4	AB008451 Canis fam
30	78	48.4	4473	6	AR080259 Sequence
31	78	48.4	4473	6	AR167390 Sequence
32	78	48.4	4473	9	HSEB2R X03363 Human c-erb
33	78	48.4	4530	6	AR202597 Sequence
34	78	48.4	4530	6	AR283481 Sequence
35	78	48.4	4530	6	AX282577 Sequence
36	78	48.4	4530	6	AX587649 Sequence
37	78	48.4	4530	6	AX644071 Sequence
38	78	48.4	4530	6	BD005474 Cellular
39	78	48.4	4530	6	I21124 Sequence 9
40	78	48.4	4530	6	I59745 Sequence 9
41	78	48.4	4530	9	HUMHER2A M1730 Human tyros
42	78	48.4	9274	6	AX060703 Sequence
43	75	46.6	528	10	AF393158 Rattus no
44	75	46.6	239732	2	AC098491 Rattus no
45	74	46.0	13450	9	AB096612 Homo sapi

ALIGNMENTS

RESULT 1


```

AX380942
LOCUS AX380942 2763 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 28 from Patent WO0212341.
ACCESSION AX380942
VERSION AX380942.1 GI:19575786
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE
1. .2763
AUTHORS Cheever,M.A. and Gheysen,D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 28 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
1. .2763
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/notes="mouse ECD-PD fusion protein cDNA"
BASE COUNT 571 a 855 c 772 g 565 t
ORIGIN
Alignment Scores:
Pred. No.: 5,57e-12 Length: 2763
Score: 153.00 Matches: 29
Percent Similarity: 96.67% Conservative: 0
Best Local Similarity: 96.67% Mismatches: 1
Query Match: 95.03% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AX380942 (1-2763)

QY 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
1918 GAACGAGGCTGCCAGCAGACGAGAGCCAGCCAGTGCAGTCTCAGACGAGGACTTA 1977
Db
QY 21 GlyProSerProMetAspSerThrPhe 30
1978 GGCCCTCCAGCCCGCATGGACAGCACCTTC 2007
Db

RESULT 2
LOCUS AX380944 2781 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 30 from Patent WO0212341.
ACCESSION AX380944
VERSION AX380944.1 GI:19575787
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE
1. .2781
AUTHORS Cheever,M.A. and Gheysen,D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 30 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
1. .2781
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/notes="mouse ECD-PD-TcP0 fusion protein cDNA"
BASE COUNT 574 a 859 c 779 g 569 t
ORIGIN
Alignment Scores:
Pred. No.: 5,61e-12 Length: 2781
Score: 153.00 Matches: 29
Percent Similarity: 96.67% Conservative: 0
Best Local Similarity: 96.67% Mismatches: 1
Query Match: 95.03% Indels: 0
DB: 6 Gaps: 0

```

```

JUNC_SEQ8_SEQ2RES991_ (1-30) x AX380944 (1-2781)

QY 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
1918 GAACGAGGCTGCCAGCAGACGAGAGCCAGCCAGTGCAGTCTCAGACGAGGACTTA 1977
Db
QY 21 GlyProSerProMetAspSerThrPhe 30
1978 GGCCCTCCAGCCCGCATGGACAGCACCTTC 2007
Db

RESULT 3
LOCUS BC027080 2083 bp mRNA linear ROD 16-APR-2003
DEFINITION Mus musculus v-erb-b2 erythroblastic leukemia viral oncogene
homolog 2, neuro/glioblastoma derived oncogene homolog (avian),
mRNA (cDNA clone MGC:38648 IMAGE:5356166), complete cds.
ACCESSION BC027080
VERSION BC027080.1 GI:20071980
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
1 (bases 1 to 2083)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshikiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Butterfield,J.S., Krzywinski,M.I., Skalska,U., Smalius,D.E.,
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2083)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 54 Row: c Column: 18
This clone was selected for full length sequencing because it

```

REFERENCE	1	Mammalia, Eutheria, Rodentia, Sciurognathia, Muridae, Mus.
AUTHORS	Spies, A.G.	
TITLE	Murine neu sequences and methods of use therefor	
JOURNAL	Patent: WO 0148205-A 1 05-JUL-2001;	
	CORIXA CORPORATION (US)	
FEATURES	Location/Qualifiers	
source	1..3771	
	/organism="Mus musculus"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10090"	
BASE COUNT	808 a 1105 c 1068 g	790 t
ORIGIN		

REFERENCE
AUTHORS Cheever, M. A. and Gheysen, D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 10 14-FEB-2002;
CORIXA CORPORATION (US); SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES Location/Qualifiers
source 1. .3955
organism="Rattus norvegicus"

```
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/note="rat Her-2/neu cDNA"
26..3799
/feature="rat Her-2/neu protein"
/codon_start=1
/protein_id="CAD28393.1"
/db_xref="GI:19575770"
/translation="MELAAWCVWGLLALLPPGIAGTQVCTGTGTMKRLPLASPETHLD
MLRHYGCGVQVGNLETVVPANASISFLQDIOEVGYMLIAHNVKRVPLQRLRV
RGTOLEFEDKVALVDNRDQDVAATPGRTPEGLRELQLRSLTEILKGLVLRGNP
QCTQDVMWNRDVRKKNQALAPVDIDNRRACPPCAPCKDNHCWGESPEDCQILTG
TIDTESMHPGRTFGASCVTTCYNYLSTEVGSCITLVCPPNNOEYVADGTQCEK
CSKPCFARGLGMEHRLGARAITSDNNVDFDGCKKIFGLSLAFLEPDPSSGIAP
LHSLRLSLEFLEITGLYLVISAMPDLSRLDSVFNLRIRGRILHDGAYSLTLQGLG
ISGLVNSLCAHGHWGPGTQVCNCSHFLRGQCEVRCVWKGPLREYVSDKRLCPCH
PECPQNSSEFCFGEADQCAAHYKDSSECVARCPGSKVPLSYMPIWKYPDEGI
QCPPIINHTSCVDLDERGCPAQRASPTVFIATVEGVLLFLVYVGVGILIKRRRQ
KIRKYMRLRLQETELVPLTPSGAMPNQAKMLKELTKRVKVLGSGAGFTVYKGI
WIPDGNVKIPVAIKVLENTSPKANKEILDEAVYVAGVSPVYVSRILGLTSTVOL
VTOLMPYKGLLDVHREHGRGLGSQDILMWCVOIAKMSYLEDLVRLHRLAARNLVK
SPNHVKTIDGLARLLIDETEYHADGCKVPKMALESILRRFTHQSDVWSTGVTV
WELTFGAKPTDGIAREIPDLLEKGERLPQPICTIDVYIMVYKMWIDSECRPRR
ELVSEFSRMARDPQRFVYIQNEDLGPSPMDSTFYRLSLEDODMDGLVDAEYLVPOQ
GFFSPDPTGRTAHRHRSSTSRSGGELTLGLEPSEEGPPRSLAPSEGAGSDVF
DGLAMGVTLQSLDPLQYSEDPTLPLPPTDGVVAPLACSPQPEVYNQSE
VOPQPLTPPEGLPPVRPAGATLERKTLSPGKNGVVKDVFATGGAVENEYLYVPRG
TASPPHSPAFSPADNLYWDQNSSEQPPSPFETPTAENPEYLGLDVPPV"
26..1990
/feature="extracellular domain (ECD) of rat Her-2/neu"
2057..3796
/feature="intracellular domain (ICD) of rat Her-2/neu"
2188..3022
/feature="kinase domain (KD) of rat Her-2/neu"
2999..3796
/feature="phosphorylation domain (PD) of rat Her-2/neu"
2999..3173
/feature="preferred portion of the phosphorylation domain
(delta PD) of rat Her-2/neu"
842 a 1147 c 1136 g 830 t

BASE COUNT      842 a 1147 c 1136 g 830 t
ORIGIN
Alignment Scores:
Pred. No.:      0.0546      Length:      3955
Score:          84.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches:  2
Query Match:    52.17%      Indels:      0
DB:             6          Gaps:          0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AX380924 (1-3955)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 2990 GTGTCATCCAGACGAGGACTTGGGCCCATCCAGCCCATGGACAGTACCTTC 3043

RESULT 7
LOCUS      121129      3955 bp      DNA      linear      PAT 07-OCT-1996
DEFINITION Sequence 14 from patent US 5518885.
ACCESSION 121129
VERSION   121129.1 GI:1601483
KEYWORDS
SOURCE    Unknown.
ORGANISM  Unknown.
REFERENCE 1 (bases 1 to 3955)
AUTHORS  Raziuddin and Sarkar, F.H.
TITLE    ERBB2 promoter binding protein in neoplastic disease
JOURNAL  Patent: US 551885-A 14 21-MAY-1996,
FEATURES  Location/Qualifiers

misc_feature
26..1990
/feature="extracellular domain (ECD) of rat Her-2/neu"
2057..3796
/feature="intracellular domain (ICD) of rat Her-2/neu"
2188..3022
/feature="kinase domain (KD) of rat Her-2/neu"
2999..3796
/feature="phosphorylation domain (PD) of rat Her-2/neu"
2999..3173
/feature="preferred portion of the phosphorylation domain
(delta PD) of rat Her-2/neu"
842 a 1147 c 1136 g 830 t

BASE COUNT      842 a 1147 c 1136 g 830 t
ORIGIN
Alignment Scores:
Pred. No.:      0.0546      Length:      3955
Score:          84.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches:  2
Query Match:    52.17%      Indels:      0
DB:             6          Gaps:          0

JUNC_SEQ8_SEQ2RES991_ (1-30) x I59750 (1-3955)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 2990 GTGTCATCCAGACGAGGACTTGGGCCCATCCAGCCCATGGACAGTACCTTC 3043

RESULT 9
LOCUS      121129      3955 bp      mRNA      linear      ROD 30-MAR-1995
DEFINITION Rat mRNA for neu oncogene (p185) encoding an epidermal growth
factor receptor-related protein.
ACCESSION X03362
VERSION   X03362.1 GI:56745
KEYWORDS  glycoprotein; kinase; neu oncogene; oncogene; transmembrane
protein; tyrosine kinase.
SOURCE    Rattus norvegicus (Norway rat)
ORGANISM  Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 3955)
AUTHORS  Bargmann, C.I., Hung, M.C. and Weinberg, R.A.
TITLE    The neu oncogene encodes an epidermal growth factor
receptor-related protein
JOURNAL  Nature 319 (6050), 226-230 (1986)
FEATURES  Location/Qualifiers
```

PUBMED REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED FEATURES source	3945311 2 (bases 1 to 3955) Lofts,F.J., Hurst,H.C., Sternberg,M.J. and Gullick,W.J. Specific short transmembrane sequences can inhibit transformation by the mutant neu growth factor receptor in vitro and in vivo Oncogene 8 (10), 2813-2820 (1993)	Location/Qualifiers 1. 3955 /organism="Rattus norvegicus" /mol_type="mRNA" /db_xref="taxon:10116" 17. .3799 /note="put. p185 precursor" /codon_start=1 /protein_id="CAA27059.1" /db_xref="GI:56746" /translation="MIIMELAANCWGFLLALLPPGIAGTQVCTGDMKRLPASPET HLMRLHLYQCGVVGNGLELTYVPANASLSFLQIOEVQGYMLIAHNOVKRVPQLRL RIVRGTLQFEDKYALAVLDNRDPQDNVAASPGRTPELRELOLRSLTEILKGVLLR GNPOLCQDMVWLKDVFKKNQLAPVDIDTNRACPPACAPACKDNHCWGSPEDCQI LTGDTTSCARCARGRLPTDCCHQCAAGCTGPKHSDCLACHFNHSGICELHCPALV CEKSKPCARVCYGLGMEHLRGARAITSDNVQEPDGCKIFGSLAFLPESDPGDPSSG IAPLRPEOLQVFELEITGYLISAWPDSLRDLSVFQNLRIIRGLIHDCGAYSLTQ GLGTHLSRLSRLGSLALIHNAHLFCVHTVPDQLEFRNPQALHLSGNRPEEDL CVSHGVCNLSAHCWGHPGPTOCVNCSHFLRGQCEVECKRWKGLPRETVSDKRL CLPPECQPNSESTFGSEADQCAACAHYKDSSCVACPCSGVKDPLSPYIWKYDE EGICQPCINTHSCVDLDERGPAEQRASPTFIATIEGVGLFLDLVVVVGILIKR RQIRKRYTRHLQETLEVEPLTPSGAMPNQAMRIILKTELKRVKVLVSGAGTGVY KGIWPDENVKIPVAIKVHRENTSPKANKEILDEAYVMAGVGSPPVYSLRILGICLTST VOLYTLMPYCCLLDHVREHREGLGSODLLNWCYQIAKMSGLIEDVRLVHRDLAARNV LVKSPNVKIDFGLARLLDIDETEHADGKVPFKWMALESILRRPFRHSDVWSYG VYVWELTFGAKYDGPAPREIPDLLEKGERLPQPICTIDVYIMVYKCMWIDSECRP RTVLESEFSRMARDPQFVVIQEDLGSPSPMDSTFYRSLLDDMDGLVDABEYLV PQCGFFPDPTPGTGAHRRHSRSTSGGELTLGLEPSEGGPPRSPDLAPSEGAGS DVFDGLAMGTGKLSLPHDSLQSLORYSEDPTLPLPETDGYVAPLACSPQPEYVN QSEYQOPPLTPEGLPVRPAGATLERPKTSLPGKNGYKDVFAFGAVENPEYLV REGTASPHSPSPAFDNLVYWDQNSSEGGPPSPFEGPTAENPEYLGIDVPV" 17. .73
sig_peptide	/note="put. signal peptide (aa -19 to -1)"	
mat_peptide	/product="put. p185 (aa 1-1241)"	
misc_feature	/note="pot. glycosylation site"	
misc_feature	/note="pot. glycosylation site"	
misc_feature	/note="pot. glycosylation site"	
misc_feature	/note="pot. glycosylation site"	
misc_feature	/note="pot. glycosylation site"	
misc_feature	/note="pot. glycosylation site"	
misc_feature	/note="pot. glycosylation site"	
misc_feature	/note="pot. phosphorylation site"	
misc_feature	/note="pot. phosphorylation site"	
misc_feature	/note="pot. phosphorylation site"	
misc_feature	/note="pot. phosphorylation site"	
misc_feature	/note="pot. phosphorylation site"	
misc_feature	/note="pot. phosphorylation site"	
BASE COUNT	842 a 1147 c 1136 g 830 t	
ORIGIN		
Alignment Scores: Pred. No.: Score:	0.0546 84.00	Length: Matches: 3955 16

Percent Similarity:
Best Local Similarity:
Query Match:
DB:

88.89%
88.89%
52.17%
10

Conservative: 0
Mismatch: 2
Indels: 0
Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x RNNEUR (1-3955)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
|||
|||||
Db 2990 GTGGTCATCCAGAACGAGGACTTGGGCCCATCCAGCCCATGGACAGTACCTTC 3043

RESULT 10
BC046811
LOCUS
DEFINITION
MUS musculus v-erb-b2 erythroblastic leukemia viral oncogene
homolog 2, neuro/glioblastoma derived oncogene homolog (avian),
mRNA (CDNA clone MGC:61178 IMAGE:5702040), complete cds.
BC046811
VERSION
BC046811.1 GI:28386210
KEYWORDS
MGC.
SOURCE
MUS musculus
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 4694)
REFERENCE
AUTHORS
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Sanchez,A., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Fay,J., Heltan,E., Kettman,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,S., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalski,U., Smalls,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
MEDLINE
PUBMED
12477932
REFERENCE
AUTHORS
Strausberg,R.
TITLE
Direct Submission
JOURNAL
Submitted (13-FEB-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,
Casavant,T., Soares,M.B.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0

misc_feature /note="assembly_fragment"
44095. .83787
/note="assembly_fragment"
83888. 96462
/note="assembly_fragment"
96563. .108784
/note="assembly_fragment"
108885. .122875
/note="assembly_fragment"
122976. .139026
/note="assembly_fragment"
139127. .159729
/note="assembly_fragment"
159830. .178173
/note="assembly_fragment"
178274. .200970
/note="assembly_fragment"
201071. .214019
/note="assembly_fragment"
vector_side:right
clone_end:77
BASE COUNT 52809 a 53348 c 53260 g 53244 t 1358 others
ORIGIN

Alignment Scores:
Pred. No.: 8.06 Length: 214019
Score: 81.50 Matches: 18
Percent Similarity: 67.86% Conservative: 1
Best Local Similarity: 64.23% Mismatches: 6
Query Match: 50.62% Indels: 3
DB: 2 Gaps: 1

JUNC_SEQ8_SEQ2RES991_ (1-30) x AC064803 (1-214019)

QY 3 GlyCysProAlaGluArgAlaSerProValThrPheGlnAsnGluAspIeuGlyPro 22
Db 53605 GCGCT------CACCTCGTGCTCCCTTCTCTAGACGAGGACTTAGGCCCC 53555
QY 23 SerSerProMetAspSerThrPhe 30
Db 53554 TCCACCCCATGGACGACCTTC 53531

RESULT 15
HAMNEU
LOCUS HAMNEU 4062 bp mRNA linear ROD 24-JAN-2003
DEFINITION Mesocricetus auratus mRNA for p-185 precursor, complete cds.
ACCESSION D16295
VERSION D16295.1 GI:493236
KEYWORDS EGF-receptor; cellular oncogene; neu-differentiation factor/herregulin; tyrosine kinase.
SOURCE Mesocricetus auratus (golden hamster)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus.

REFERENCE 1 (bases 1 to 4062)
AUTHORS Nakamura,T., Ushijima,T., Ishizaka,Y., Nagao,M., Arai,M., Yamazaki,Y. and Ishikawa,T.
TITLE Cloning and activation of the Syrian hamster neu proto-oncogene
JOURNAL Gene 140 (2), 251-255 (1994)
MEDLINE 94193007
PUBMED 7908275
REFERENCE 2 (bases 1 to 4062)
AUTHORS Nakamura,T.
TITLE Direct Submission
JOURNAL Submitted (19-MAY-1993) Takuro Nakamura, Faculty of Medicine, University of Tokyo, Department of Pathology; 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan (Tel:81-3-3812-2111(ex.3356), Fax:81-3-3815-8379)
FEATURES Location/Qualifiers
source 1..4062
/organism="Mesocricetus auratus"
/mol_type="mRNA"

/isolate="animal 14"
/db_xref="taxon:10036"
/clone="pSHN"
/cell_line="14-2"
/cell_type="neoplastic Schwann cell"
/tissue_type="peripheral nerve"
/clone_lib="NIH3T3"
/note="common name: Syrian golden hamster"
1. 4062
/gene="neu"
33. .3797
/gene="neu"
/codon_start=1
/product="p-185 precursor"
/protein_id="BAA03801.1"
/db_xref="GI:747595"
/translation="MELAAWCGWLLALLSPGASGTOVCTGCTDNKMLRLPASPEHLD
IVRHLYOGCVQVQNNLELYLPANATLSFLQDIOEVGYMLIAHSQVRHVPQLRLRV
RGTLQFEDKYALAVLNDRLDPLDVTATGRTPEGLRELQSLRSLTEILKGVLLIRNPQ
LCYQDTVLWKDVFRRNQLAPVDITNRSRACPCAPACAKONHGWGASPCQCLTGT
IAPRAVPAARARLPTDCHEQCAAGCTGPKHSDCLACILHNSHGICELCHCALVTYNT
DTFESMPNPEGRTYFGASCVTTCYNYLSTEVGCTLVCLPLNNOEVTADCTQCECK
SKSCARVCYGLMEHLRGARAITSANIOEFAGCKIFGSLAFLPESFDGNSSGAPL
TPEQLQVFETLEEITGYLISAWPDSLDHLSVFQNLRVIRGRVLDHGDGAYSLAQGLGI
RWGLRLSRELGSGLVLIHRNTHLCFVHTVPDQDLFRNPHQALLHSGNPSEEGCLKD
FACYPLCAHGWGPGTQCYNVCSHFRLGQCEKRCRVKGLPREYVNGKCLPCHPE
COPONSTETCTGSEADQCTACPHYKDSFPCVAPCPGVKPDLSYMPWKYDDEGMCQ
PCPINCHSCVDLDERGCPAQSRAPATSIATVVGILLFLVIGVVGILLKRRQKI
RKYTMRLQLTELVEPLTPSGAMPQAMRILKTELKRVKVLGSGAFGTYVYGIWI
PDGENKIPVAIKLRENTSPKANKEILDEAYVMAGLSYVSRLLGLTCLTSTVOLVT
QLMPYGCILLDHVREHRLGSDLLNMVCQIAKMSYLEDVRLVHRDLAARNLVKSP
NHVKTDFGLARLLIDETEHADGKVPIKWLALIESILRRRFTHOSDWSYGVTVME
LMTFGAPYDGIIPAREIPDLLEKGERLPQPPCTIDVYIMVIMVKWIMIDSECRPREL
VFSPRWARDQRFVWVQNEDLGSPSSPLDSTFYASLLEDDDMGLVDAEELVYPOQFG
FFPDPAAGSTAHRHRSSTSGGGELTGMEPSEEPPEPRSLAPSEAGSVDVEG
ELHGATKGPOSISPRDLSPQRYSEDPTLPTDTGTVAPLACSPQPEYVNOPEVR
PQPLLTPEGPLPVRPAGATLERPKTLSPGKNVYKDVFTFGGAVENPEYLVPRGSA
SQPHPALCPAFDNLVYWDQPSERGSPPNTFEGTPTAENPEYLGLDVVPV"
33. .86
/gene="neu"
/note="putative"
87. .3794
/gene="neu"
/product="p-185"
516. .1988
/gene="neu"
/note="extracellular domain"
1989. .2057
/gene="neu"
/note="putative
transmembrane domain"
2058. .3794
/gene="neu"
/note="cytoplasmic domain"
BASE COUNT 845 a 1209 c 1182 g 826 t
ORIGIN

Alignment Scores:
Pred. No.: 0.15 Length: 4062
Score: 81.00 Matches: 15
Percent Similarity: 88.89% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 2
Query Match: 50.31% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x HAMNEU (1-4062)

QY 13 ValThrPheGlnAsnGluAspIeuGlyProSerSerProMetAspSerThrPhe 30
Db 2991 GTGGTCATCCAGAAATGAGACTTGGGCCCTTCCAGCCCTTGGACAGATACCTTT 3044

Search completed: October 16, 2003, 11:01:42

Job time : 465.208 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 234.794 Seconds
(without alignments)
3105.423 Million cell updates/sec

Title: JUNC_SEQ3_SEQ5
Perfect score: 159
Sequence: 1 DKGCPAQRASPLTSONEDIGPASPDLSTF 30

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool_P/HOLLERAN480/runat_15102003_131913_20548/app_query.fasta_1.4685
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 @CGN_1_1_12645 @runat_15102003_131913_20548 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	85.5	53.8	618	28	AZ699313 RPCI-23-2
2	85	53.5	422	9	AW606421 RC3-HT037
3	85	53.5	439	10	BE157169 RC3-HT037
4	85	53.5	944	14	CD558559 AGENCOURT
5	81	50.9	458	10	BE157103 RC3-HT037
6	80	50.3	378	10	BF869854 IL3-EF011
7	80	50.3	411	2	BSM069572 Homo sapi
8	80	50.3	433	10	BF762317 IL2-CS004
9	80	50.3	450	12	BG988031 MR2-HT116
10	80	50.3	466	12	BM720186 UI-E-E00-
11	80	50.3	546	12	BI194790 602948150
12	80	50.3	547	9	AA443351 zw85b05.i
13	80	50.3	561	12	BM720098 UI-E-E00-
14	80	50.3	607	12	BM083444 imagegc_6
15	80	50.3	613	9	AW410534 fh06h06.x
16	80	50.3	666	13	BX113765 BX113765
17	80	50.3	667	12	BM048624 603623984
18	80	50.3	674	12	BM758528 K-EST0038
19	80	50.3	741	10	BE615590 601278973
20	80	50.3	881	9	AI906106 RC-BT105-
21	80	50.3	886	13	BO882315 AGENCOURT
22	80	50.3	893	14	CA487631 AGENCOURT
23	80	50.3	905	13	BU541491 AGENCOURT
24	80	50.3	909	10	BG289457 602381495
25	80	50.3	922	14	CA453990 AGENCOURT
26	80	50.3	1009	13	BO647343 AGENCOURT
27	79	49.7	278	10	BF923905 OVA-NT025
28	79	49.7	294	10	BF757857 CMA-CT057
29	79	49.7	337	10	BE091682 IL2-BT073
30	79	49.7	348	9	AI909847 QV-BT225-
31	79	49.7	378	14	CB266492 1005398.H
32	79	49.7	423	12	BM703962 UI-E-CK1-
33	79	49.7	464	2	BSM067221 Homo sapi
34	79	49.7	496	12	BM795191 K-EST0076
35	79	49.7	505	12	BM854421 K-EST0138
36	79	49.7	531	12	BM830170 K-EST0103
37	79	49.7	536	14	CB129403 K-EST0179
38	79	49.7	547	12	BM787824 K-EST0066
39	79	49.7	567	2	BSM077666 Homo sapi
40	79	49.7	574	12	BM829991 K-EST0103
41	79	49.7	630	2	BSM073298 Homo sapi
42	79	49.7	691	14	CB853376 UI-CF-FNO
43	79	49.7	767	10	BF240297 601905830
44	79	49.7	849	14	CA489799 AGENCOURT
45	79	49.7	852	14	CD516283 AGENCOURT

ALIGNMENTS

RESULT 1
AZ699313
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

AZ699313
RPCI-23-240N16.TJ RPCI-23
Mus musculus genomic clone
, genomic survey sequence.
GI:12419623
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 618)

618 bp
DNA
linear
GSS 24-JAN-2001
GSS 24-JAN-2001

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-RC3-HT0371-250>)
200-016-all6t3-2000-02-256t4-1)
Seq primer: puc 18 forward
High quality sequence stop: 439.

FEATURES
source
1. .439
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0371"
/notes="Organ: head,neck; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 89 a 139 c 132 g 79 t

ORIGIN

Alignment Scores:
Pred. No.: 0.262 Length: 439
Score: 85.00 Matches: 23
Percent Similarity: 46.00% Conservative: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match: 53.46% Indels: 24
DB: 10 Gaps: 2

JUNC_SEQ3_SEQ5 (1-30) x BE157169 (1-439)

QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
|||||
Db 167 CCGCAGAGGGTGGGAAGGAGAGATGAGTCCAGTATGCCAGGCCCTCACGGAAGGCTGC 108
QY 16 -----GlnAsnGluAspLeu 20
|||||
Db 107 ATGCTGGGCTGGGAGGGCCACCACCTGCTCTCTCTCCATCCACAGATGAGGACTTG 48
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db 47 GGCCCGCAGTGCCTTGGACAGCACCTTC 18

RESULT 4

LOCUS CD558559 944 bp mRNA linear EST 11-JUN-2003

DEFINITION AGENCOURT_14477527 NIH_MGC_181 Homo sapiens cDNA clone

ACCESSION IMAGE:30396701 5', mRNA sequence.

VERSION CD558559

KEYWORDS CD558559.1 GI:31584627

SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 944)

TITLE NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
CDNA Library Prepared by: Dr. Michael Brownstein
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>

Plate: NDAM478 row: n column: 06
High quality sequence start: 19
High quality sequence stop: 537.

FEATURES

source

1. .944
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="IMAGE:30396701"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="NIH_MGC_181"
/clone_lib="NIH_MGC_181"
/notes="Vector: pCMV-SPORT6.1; Site.1: NotI; Site.2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.42 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."

BASE COUNT 150 a 269 c 366 g 158 t

ORIGIN

Alignment Scores:
Pred. No.: 0.682 Length: 944
Score: 85.00 Matches: 23
Percent Similarity: 46.00% Conservative: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match: 53.46% Indels: 24
DB: 14 Gaps: 2

JUNC_SEQ3_SEQ5 (1-30) x CD558559 (1-944)

QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
|||||
Db 193 CCGCAGAGGGTGGGAAGGAGAGATGAGTCCAGTATGCCAGGCCCTCACGGAAGGCTGC 252
QY 16 -----GlnAsnGluAspLeu 20
|||||
Db 253 ATGCTGGGCTGGGAGGGCCACCACCTGCTCTCTCTCCATCCACAGATGAGGACTTG 312
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db 313 GGCCCGCAGTGCCTTGGACAGCACCTTC 342

RESULT 5

LOCUS BE157103 458 bp mRNA linear EST 21-JUN-2000

DEFINITION RC3-HT0371-180200-015-h07 HT0371 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE157103

VERSION BE157103.1 GI:8619824

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 458)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IH3&t2=IL3-ET0114-251000-317-B07&t3=2000-10-25&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 177.
Location/Qualifiers
1..378
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="ET0114"
/note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESPEC PCR (U.S. Letters Patent application No. 196.716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 76 a 119 c 110 g 73 t
ORIGIN
Alignment Scores:
Pred. No.: 0.944 Length: 378
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 10 Gaps: 0
JUNC_SEQ3_SEQ5 (1-30) x BF869854 (1-378)
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 237 GACAAGGGTGCCCCCGCAGCAGAGAGCCGCGCCCTTGACGTC 193
RESULT 7
HSM069572
ID HSM069572 standard; RNA; EST; 411 BP.
XX
AC BX479682;
XX
SV BX479682.1
XX
DT 09-MAY-2003 (Rel. 75, Created)
DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)
XX
DE Homo sapiens mRNA; EST DKFZp686K212l3_r1 (from clone DKFZp686K212l3)
XX
EST; expressed sequence tag.
XX
KW Homo sapiens (human)
XX OS
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
[1]
RN 1-411
RP Bloeker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RA Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.
RL MIPB, Inqolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
RL

```

XX CC This is the 5' sequence of the clone insert
CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC sequenced by GBF (National Research Centre for Biotechnology
CC Ltd., Braunschweig/Germany) within the cDNA sequencing
CC consortium of the German Genome Project.
CC No s1 sequence available.
CC This clone (DKFZp686K21213) is available at the RZPD in Berlin.
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX FH
XX FH
XX FH
XX FT
XX FT
XX FT
XX FT
XX FT
XX FT
XX FT
XX FT
XX SQ Sequence 411 BP; 70 A; 126 C; 125 G; 90 T; 0 other;

Alignment Scores:
Pred. No.: 1.05 Length: 411
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x HSM069572 (1-411)

QY 1 AsPLysGLyCySProAlaGLuGlnArgAlaSerProLeuThrSer 15
|||||
DB 292 GACAAGGCTGCCCGCCGAGCAGAGACCCCTCTGACGTCC 336

RESULT 8
BF762317/c 433 bp mRNA linear EST 12-JAN-2001
LOCUS
DEFINITION IL2-CS0049-181000-202-A06 CS0049 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF762317
VERSION BF762317.1 GI:12110217
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 433)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-HT1160-
110101-005-a06&t3=2001-01-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 39
High quality sequence stop: 325.

```

```

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL2&t2=IL2-CS0049-
181000-202-A06&t3=2000-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 387.
Location/Qualifiers
1..433
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CS0049"
/note="Organ: colon_est; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 86 a 149 c 108 g 90 t
ORIGIN

Alignment Scores:
Pred. No.: 1.12 Length: 433
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ5 (1-30) x BF762317 (1-433)

QY 1 AsPLysGLyCySProAlaGLuGlnArgAlaSerProLeuThrSer 15
|||||
DB 367 GACAAGGCTGCCCGCCGAGCAGAGACCCCTCTGACGTCC 323

RESULT 9
BG988031 450 bp mRNA linear EST 13-JUN-2001
LOCUS
DEFINITION MR2-HT1160-110101-005-a06 HT1160 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG988031
VERSION BG988031.1 GI:14392101
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 450)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-HT1160-
110101-005-a06&t3=2001-01-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 39
High quality sequence stop: 325.

```

```

FEATURES
source
  Location/Qualifiers
    1..450
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /dev_stage="Adult"
      /clone_lib="HT1160"
      /note="Organ: head,neck; Vector: puc18; Site_1: SmaI;
      Site_2: SmaI; A mini-library was made by cloning products
      derived from ORESTES PCR (U.S. Letters Patent application
      No. 196,716 - Ludwig Institute for Cancer Research)
      profiles into the pUC 18 vector. Reverse transcription of
      tissue mRNA and cDNA amplification were performed under
      low stringency conditions."
BASE COUNT      84 a 127 c 150 g 89 t
ORIGIN
Alignment Scores:
Pred. No.:      1.17      Length:      450
Score:          80.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     50.31%      Indels:      0
DB:              12      Gaps:         0

JUNC_SEQ3_SEQ5 (1-30) x BG988031 (1-450)

QY      1 AsPLySGlYcysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db      259 GACAAGGGCTGCCCTGCCAGCAGAGAGCCAGCCCTCTGACGTC 303

RESULT 10
BM720186
LOCUS      BM720186      466 bp      mRNA      linear      EST 01-MAR-2002
DEFINITION UI-E-E00-ahz-c-08-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone
            UI-E-E00-ahz-c-08-0-UI 5', mRNA sequence.
ACCESSION  BM720186
VERSION     BM720186.1 GI:19039265
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 466)
AUTHORS     Bonaudo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
PUBMED      8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
            Location/Qualifiers
              1..466
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="UI-E-E00-ahz-c-08-0-UI"
                /tissue_type="fetal eye"
                /dev_stage="fetal"
                /lab_host="DH10B (Life Technologies) (T1 phage resistant)"

FEATURES
source
  Location/Qualifiers
    1..546
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:5091396"
      /tissue_type="epithelioid carcinoma cell line"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH_MGC_42"
      /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
      Site_2: EcoRI; cDNA made by oligo-dT priming.
      Directionally cloned into EcoRI/XhoI sites using the

```

```

/clone_lib="UI-E-E00"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-E00 is a cDNA library containing the following
tissue(s): fetal eye. The library was constructed
according to Bonaudo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor. digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGCGTATACC. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
BASE COUNT      105 a 114 c 151 g 93 t 3 others
ORIGIN
Alignment Scores:
Pred. No.:      1.23      Length:      466
Score:          80.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     50.31%      Indels:      0
DB:              12      Gaps:         0

JUNC_SEQ3_SEQ5 (1-30) x BM720186 (1-466)

QY      1 AsPLySGlYcysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db      18 GACAAGGGCTGCCCTGCCAGCAGAGAGCCAGCCCTCTGACGTC 62

RESULT 11
BM720186
LOCUS      BM720186      546 bp      mRNA      linear      EST 10-JUL-2001
DEFINITION 602948150F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5091396 5',
            mRNA sequence.
ACCESSION  BM720186
VERSION     BM720186.1 GI:14649810
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 546)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLML at:
            http://image.llnl.gov
            Plate: LLCMI855 row: c column: 13
            High quality sequence start: 10
            High quality sequence stop: 495.
FEATURES
source
  Location/Qualifiers
    1..546
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:5091396"
      /tissue_type="epithelioid carcinoma cell line"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH_MGC_42"
      /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
      Site_2: EcoRI; cDNA made by oligo-dT priming.
      Directionally cloned into EcoRI/XhoI sites using the

```

following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library. |"

BASE COUNT 110 a 147 c 183 g 106 t

ORIGIN

Alignment Scores:
Pred. No.: 1.49 Length: 546
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x B1194790 (1-546)

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

|||||
Db 223 GACAAGGGCTGCCCCCGGAGCAGAGAGCCGCTCTGACGTC 267

RESULT 12

AA443351

LOCUS

DEFINITION AA443351 547 bp mRNA linear EST 03-JUN-1997
IMAGE:783729 5' similar to gb:U11730 ERBB-2 RECEPTOR
PROTEIN-TYROSINE KINASE PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION

AA443351

VERSION

AA443351.1 GI:2156026

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 547)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie,
T., Waterston, R., and Wilson, R.
WashU-Merck EST Project 1997

TITLE

JOURNAL

COMMENT

Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 360.

FEATURES

source

1..547
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5981347"
/db_xref="taxon:9606"
/clone="IMAGE:783729"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares total fetus Nb2HF8.9w"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site: Not I; Site:2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAATGGGAGCGCGCTTAATTTTTTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 115 a 151 c 172 g 108 t 1 others

ORIGIN

Alignment Scores:
Pred. No.: 1.5 Length: 547
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 9 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x AA443351 (1-547)

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

|||||
Db 174 GACAAGGGCTGCCCCCGGAGCAGAGAGCCGCTCTGACGTC 218

RESULT 13

BM720098

LOCUS

DEFINITION BM720098 561 bp mRNA linear EST 01-MAR-2002
UI-E-E00-ahz-c-05-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone
UI-E-E00-ahz-c-05-0-UI 5', mRNA sequence.

ACCESSION

BM720098

VERSION

BM720098.1 GI:19039119

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 561)

AUTHORS

Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

1..561

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-E-E00-ahz-c-05-0-UI"

/tissue_type="fetal eye"

/dev_stage="fetal"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-E00"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site:1: Eco RI; Site:2: Not I;

UI-E-E00 is a cDNA library containing the following

tissue(s): fetal eye. The library was constructed

according to Bonaldo, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an Eco RI adaptor, digested

with Not I, and cloned directionally into pT7T3-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tag for this library is GCGGTATACC. This library

NOTES

JUNC_SEQ3_SEQ5 (1-30) x AW4

STANDARD

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||
Db 557 GACAAGGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGACGTCC 601

Search completed: October 16, 2003, 17:04:03
Job time : 235.794 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 11:08:07 ; Search time 31.0111 Seconds
(without alignments)
2540.503 Million cell updates/sec

Title: JUNC_SEQ3_SEQ5

Perfect score: 159

Sequence: 1 DKGCPARQASPLTSQNEGLGASPLDSTF 30

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q/cgn2_1/USPFO_spool_P/HOLLERAN480/runat_15102003_131915_20662/app_query.fasta_1.4685
-DB=Published Applications_NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=HOLLERAN480 -CGN_1_1_1_397 -runat_15102003_131915_20662
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA :
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq :
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq :
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq :
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq :
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq :
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq :
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq :
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq :
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq :
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq :
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq :
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq :
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq :
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq :
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq :
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq :
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq :

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	80	50.3	201 13	US-10-109-213-3 Sequence 3, Appli

2	80	50.3	3765	12	US-10-207-498-5	Sequence 5, Appli
3	80	50.3	3768	9	US-09-811-123-8	Sequence 8, Appli
4	80	50.3	3768	9	US-09-811-115-2	Sequence 2, Appli
5	80	50.3	3768	10	US-09-854-356-9	Sequence 9, Appli
6	80	50.3	3768	10	US-09-930-125-1	Sequence 1, Appli
7	80	50.3	3768	12	US-10-313-644-1	Sequence 5, Appli
8	80	50.3	4473	11	US-09-441-411-5	Sequence 81, Appli
9	80	50.3	4473	12	US-10-101-510-81	Sequence 32, Appli
10	80	50.3	4473	14	US-10-146-473-32	Sequence 44, Appli
11	80	50.3	4473	14	US-10-207-655-44	Sequence 11, Appli
12	80	50.3	4530	10	US-09-877-177-11	Sequence 119, App
13	80	50.3	4530	12	US-10-007-926A-119	Sequence 124, App
14	80	50.3	4530	12	US-10-101-510-124	Sequence 1, Appli
15	80	50.3	4530	12	US-10-338-730-1	Sequence 125, App
16	80	50.3	4530	14	US-10-177-293-125	Sequence 1, Appli
17	80	50.3	4543	10	US-09-769-508-1	Sequence 70, Appli
18	80	50.3	4606	12	US-09-971-392-70	Sequence 10896, A
19	80	50.3	4642	14	US-10-198-846-10896	Sequence 7, Appli
20	80	50.3	9274	9	US-09-811-123-7	Sequence 165, App
21	80	50.3	9274	9	US-09-811-115-1	Sequence 14, Appli
22	79	49.7	1115	14	US-10-102-806-165	Sequence 6, Appli
23	79	49.7	1713	12	US-10-378-393-14	Sequence 4, Appli
24	79	49.7	1755	10	US-09-930-125-6	Sequence 7, Appli
25	79	49.7	1767	10	US-09-930-125-4	Sequence 5, Appli
26	79	49.7	1773	10	US-09-930-125-7	Sequence 10, Appli
27	79	49.7	1806	10	US-09-930-125-5	Sequence 117, App
28	79	49.7	2411	12	US-10-378-393-10	Sequence 117, App
29	74	46.5	3771	10	US-09-854-356-11	Sequence 1540, Ap
30	74	46.5	3955	10	US-09-870-759-117	Sequence 1332, Ap
31	74	46.5	3955	10	US-09-854-356-10	Sequence 1, Appli
32	74	46.5	3955	12	US-09-751-708A-117	Sequence 13, Appli
33	62	39.0	14427	14	US-10-156-761-1540	Sequence 28706, A
34	62	39.0	23432	9	US-09-764-869-1332	Sequence 11498, A
35	62	39.0	23432	14	US-10-091-504-1332	Sequence 25, Appli
36	62	39.0	9025608	14	US-10-156-761-1	Sequence 27, Appli
37	59	37.1	11182	13	US-10-094-989-3	Sequence 23, Appli
38	58.5	36.8	633	11	US-09-788-188-12	
39	58.5	36.8	634	8	US-08-450-842-1	
40	58.5	36.8	1404	14	US-10-155-785-13	
41	57.5	36.2	363	11	US-09-918-995-28706	
42	57	35.8	297	10	US-09-783-590-11498	
43	57	35.8	1938	12	US-10-422-264-25	
44	57	35.8	2316	12	US-10-422-264-27	
45	57	35.8	2604	12	US-10-422-264-23	

ALIGNMENTS

RESULT 1

US-10-109-213-3
; Sequence 3, Application US/10109213
; Publication No. US20020186670A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; TITLE OF INVENTION: Identification Of Disease Predictive Nucleic Acids
; FILE REFERENCE: IBIS0009
; CURRENT APPLICATION NUMBER: US/10/109,213
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US/09/200,355
; PRIOR FILING DATE: 1998-11-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 201
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-109-213-3

Alignment Scores:	0.000724	Length:	201
Pred. No.:	80.00	Matches:	15
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%		

Query Match: 50.31% Indels: 0
DB: 13 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-10-109-213-3 (1-201)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 116 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCUCUGACGTC 160

RESULT 2

US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREGULIN AND HER3
; FILE REFERENCE: 30448.103-US-U1
; CURRENT APPLICATION NUMBER: US/10/207,498
; CURRENT FILING DATE: 2002-07-29
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5

Alignment Scores:
Pred. No.: 0.0177 Length: 3765
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-10-207-498-5 (1-3765)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTC 1959

RESULT 3

US-09-811-123-8
; Sequence 8, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ErbB
; TITLE OF INVENTION: ANTIBODY-MAYTANSINOID CONJUGATES
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/602,530
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-811-123-8

Alignment Scores:
Pred. No.: 0.0177 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 9 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-09-811-123-8 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTC 1959

RESULT 4

US-09-811-115-2
; Sequence 2, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred. No.: 0.0177 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 9 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-09-811-115-2 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTC 1959

RESULT 5

US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768

JUNC_SEQ3_SEQ5 (1-30) x US-09-930-125-1 (1-3768)

JUNC SE03 SE05 (1-30) x US-09-441-411-5 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGGCTGCCCGCCGAGAGAGAGAGCCAGCCCTCTGACGTCC 2133

RESULT 9

US-10-101-510-81
; Sequence 81, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-81

Alignment Scores:
Pred. No.: 0.0214 Length: 4473
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-10-101-510-81 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGGCTGCCCGCCGAGAGAGAGCCAGCCCTCTGACGTCC 2133

RESULT 10

US-10-146-473-32
; Sequence 32, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-473-32

Alignment Scores:
Pred. No.: 0.0214 Length: 4473
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 14 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-10-146-473-32 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGGCTGCCCGCCGAGAGAGAGCCAGCCCTCTGACGTCC 2133

RESULT 11

US-10-207-655-44
; Sequence 44, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-207-655-44

Alignment Scores:
Pred. No.: 0.0214 Length: 4473
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 14 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-10-207-655-44 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGGCTGCCCGCCGAGAGAGAGCCAGCCCTCTGACGTCC 2133

RESULT 12

US-09-877-177-11
; Sequence 11, Application US/09877177
; Publication No. US20020192652A1
; GENERAL INFORMATION:
; APPLICANT: Peter V. Danenberg et al.
; TITLE OF INVENTION: Method of determining Epidermal Growth
; FACTOR RECEPTOR AND HER2-NEW GENE EXPRESSION
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177-11

Alignment Scores:
Pred. No.: 0.0217 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-09-877-177-11 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2065 GACAAGGGCTGCCCGCCGAGAGAGAGCCAGCCCTCTGACGTCC 2109

RESULT 13

US-10-007-926A-119

```

Qy 1 AspYsGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
   |||||
Db 2065 GACAAGGGCTGCCCGCGGAGCAGAGAGCCGCTCTGACGTC 2109

RESULT 15
US-10-338-730-1
; Sequence 1, Application US/10338730
; Publication No. US20030147905A1
; GENERAL INFORMATION:
; APPLICANT: Genzyme Corporation
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS
; FILE REFERENCE: 5017C
; CURRENT APPLICATION NUMBER: US/10/338,730
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: US 09/527,487
; PRIOR FILING DATE: 2002-03-16
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
; OTHER INFORMATION:
US-10-338-730-1

Alignment Scores:
Pred. No.: 0.0217 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-10-338-730-1 (1-4530)

Qy 1 AspYsGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
   |||||
Db 2065 GACAAGGGCTGCCCGCGGAGCAGAGAGCCGCTCTGACGTC 2109

Search completed: October 17, 2003, 03:53:03
Job time : 39.0111 secs

```


THIS PAGE BLANK (USPTO)


```
;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3

Alignment Scores:
Pred. No.: 0.0053 Length: 2385
Score: 84.00 Matches: 21
Percent Similarity: 70.00% Conservative: 0
Best Local Similarity: 70.00% Mismatches: 5
Query Match: 52.83% Indels: 4
DB: 2 Gaps: 2

JUNC_SEQ3_SEQ5 (1-30) x US-09-146-283-3 (1-2385)

QY 1 ASPLYSGLYCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGCTGCCCCCGCAGAGAGCCCTCTGACGTCCTCGAG----- 1975

QY 21 GlyProAla---SerProLeuAspSerThr 29
Db 1976 GCACCGCGCGCTCGCCAGCCAGCCAGCACA 2005

RESULT 2
US-08-579-823A-3
; Sequence 3, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-146-283-3

Alignment Scores:
Pred. No.: 0.0053 Length: 2385
Score: 84.00 Matches: 21
Percent Similarity: 70.00% Conservative: 0
Best Local Similarity: 70.00% Mismatches: 5
Query Match: 52.83% Indels: 4
DB: 2 Gaps: 2

JUNC_SEQ3_SEQ5 (1-30) x US-08-579-823A-3 (1-2385)

QY 1 ASPLYSGLYCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGCTGCCCCCGCAGAGAGCCCTCTGACGTCCTCGAG----- 1975

QY 21 GlyProAla---SerProLeuAspSerThr 29
Db 1976 GCACCGCGCGCTCGCCAGCCAGCCAGCACA 2005

RESULT 3
US-09-344-195-3
; Sequence 3, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-344-195-3

Alignment Scores:
```

Pred. No.: 0.0053 Length: 2385
Score: 84.00 Matches: 21
Percent Similarity: 70.00% Conservative: 0
Best Local Similarity: 70.00% Mismatches: 5
Query Match: 52.83% Indels: 4
DB: 3 Gaps: 2

JUNC_SEQ3_SEQ5 (1-30) x US-09-344-195-3 (1-2385)

Qy 1 AsplysGlyCysProAlaGluInArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
|||||
Db 1925 GACAAAGGCTGCCCGCGAGAGAGAGCCAGCCCTCTGAGCTCTCGAG----- 1975
Qy 21 GlyProAla---SerProLeuAspSerThr 29
|||||
Db 1976 GCACCGCGCTCGCCGAGCCAGCCAGCACA 2005

RESULT 4

US-08-776-251-3
; Sequence 3, Application US/08776251
; Patent No. 6025340
; GENERAL INFORMATION:
; APPLICANT: Springer, Caroline J
; APPLICANT: Marais, Richard
; TITLE OF INVENTION: Surface expression of enzyme in gene directed prodrug therapy
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6025340th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,251
; FILING DATE: 31-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/01782
; FILING DATE: 27-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9415167.7
; FILING DATE: 27-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 620-20
; INFORMATION FOR SEQ ID NO: 3:
; LENGTH: 153 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-776-251-3

Alignment Scores:
Pred. No.: 0.000654 Length: 153
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-08-776-251-3 (1-153)

Qy 1 AsplysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 10 GACAAAGGCTGCCCGCGAGAGAGAGCCAGCCCTCTGAGCTCC 54

RESULT 5
US-09-200-355-3
; Sequence 3, Application US/09200355
; Patent No. 6451524
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; TITLE OF INVENTION: Identification of Disease Predictive Nucleic Acids
; FILE REFERENCE: IBIS0009
; CURRENT APPLICATION NUMBER: US/09/200,355
; CURRENT FILING DATE: 1998-11-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 201
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-200-355-3

Alignment Scores:
Pred. No.: 0.000922 Length: 201
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-09-200-355-3 (1-201)

Qy 1 AsplysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 116 GACAAAGGCGGCCCGCGAGAGAGAGCCAGCCCGCUGAGCGGCC 160

RESULT 6

US-08-776-251-10
; Sequence 10, Application US/08776251
; Patent No. 6025340
; GENERAL INFORMATION:
; APPLICANT: Springer, Caroline J
; APPLICANT: Marais, Richard
; TITLE OF INVENTION: Surface expression of enzyme in gene directed prodrug ther
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6025340th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,251
; FILING DATE: 31-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/01782
; FILING DATE: 27-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9415167.7
; FILING DATE: 27-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 620-20
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-776-251-10

```

Alignment Scores:
Pred. No.: 0.00536 Length: 816
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-08-776-251-10 (1-816)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 93 GACAAGGGCTGCCCGCCGAGCAGAGACGCCGCTCTGACGTC 137
|||||

RESULT 7
US-08-625-101-1
; Sequence 1, Application US/08625101
; Patent No. 5869445
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 0.0366 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-08-625-101-1 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||

```

```

Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGACGCCGCTCTGACGTC 1959

RESULT 8
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3768
; OTHER INFORMATION: /note= "product = "cerB-b2""
US-08-356-786-1

Alignment Scores:
Pred. No.: 0.0366 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-08-356-786-1 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGACGCCGCTCTGACGTC 1959
|||||

RESULT 9
US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters

```



```
;
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPlastic DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-645-865-9

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-08-645-865-9 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2065 GACAAGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTC 2109

RESULT 13
US-09-167-322-4
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
```

```
;
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-09-167-322-4 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2065 GACAAGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTC 2109

RESULT 14
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
US-09-527-487-1

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-09-527-487-1 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2065 GACAAGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTC 2109

RESULT 15
US-09-877-177A-11
; Sequence 11, Application US/09877177A
```

```

; Patent No. 6582919
;
; GENERAL INFORMATION:
;
; APPLICANT: K. Danenberg
;
; TITLE OF INVENTION: Method of determining Epidermal Growth
;
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
;
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
;
; FILE REFERENCE: 11220/120
;
; CURRENT APPLICATION NUMBER: US/09/877,177A
;
; CURRENT FILING DATE: 2001-06-11
;
; NUMBER OF SEQ ID NOS: 11
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 11
;
; LENGTH: 4530
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; US-09-877-177A-11

```

Alignment Scores:		
Pred. No.:	0.0462	Length: 4530
Score:	80.00	Matches: 15
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	50.31%	Indels: 0
DB:	4	Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-09-877-177A-11 (1-4530)

OY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 |||||
Db 2065 GACAAGGGTGTCCCGCCGACAGAGACGACGCCCTGTACGTCC 2109

Search completed: October 16, 2003, 17:16:18
Job time : 14.6049 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:06:28 : Search time 28.4941 Seconds
(without alignments)
2842.104 Million cell updates/sec

Title: JUNC_SEQ3_SEQ5
Perfect score: 159
Sequence: 1 DKCPAEQRASPLTSQNDLGPASPLDSTF 30

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_P/HOLLERAN480/runat_15102003_131912_20526/app_query.fasta_1.4685
-DB=N_Geneseq_19Jun03 -QFWT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptlo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 @CGN_1_1758 @runat_15102003_131912_20526 -NCFU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	144	90.6	2763	24	ABA92252	Mouse Her-2/neu ex
2	144	90.6	2781	24	ABA92253	Her-2/neu extracel
3	84	52.8	2385	18	AA72723	Her2-GM-CSF immuno
4	80	50.3	69	21	AAA14815	DNA encoding a wil
5	80	50.3	201	19	AA62125	Her2 normal form t
6	80	50.3	2125	19	AAV21727	Humanised vector p
7	80	50.3	2871	21	AA250586	DC8scFv-erbB2EC fu
8	80	50.3	3600	21	AA889736	Human HER-2/neu co
9	80	50.3	3678	24	ABR66207	cDNA encoding huma
10	80	50.3	3768	17	AA740739	HER-2/neu oncogene
11	80	50.3	3768	20	AA01912	Human HER-2/neu on
12	80	50.3	3768	21	AAA09455	Human heregulin 2
13	80	50.3	3768	22	AAH23392	Human HER-2/neu pr
14	80	50.3	3768	24	AB235744	Human ERBB2 polynu
15	80	50.3	3768	24	ABX09987	Human ERBB2 DNA fr
16	80	50.3	3768	24	AA043935	Human HER-2 cDNA.
17	80	50.3	3768	24	AA043986	Human ER2 antigen
18	80	50.3	3768	24	ABV78168	Human ERBB2 DNA SE
19	80	50.3	3768	24	AA032743	Human Her-2/neu pr
20	80	50.3	3768	24	ABA92250	Human Her-2/neu CD
21	80	50.3	3768	24	ABK10730	Human Her-2/neu DN
22	80	50.3	3768	24	ABU91709	Human polynucleoti
23	80	50.3	3768	24	ABK14058	Human HER2 (ErbB2)
24	80	50.3	4299	14	AA046083	Sequence encoding
25	80	50.3	4472	21	AAA14812	cDNA encoding the
26	80	50.3	4473	19	ABQ76220	Human tumour anti
27	80	50.3	4473	20	AA231071	HER-2 nucleic acid
28	80	50.3	4473	24	AB234969	Human gene express
29	80	50.3	4473	24	AA038904	Human Her-2 DNA.
30	80	50.3	4530	16	AA01585	Her-2/neu (ERBB2)/c
31	80	50.3	4530	21	AA71253	Human ER2 gene.
32	80	50.3	4530	21	AA260815	Nucleotide sequenc
33	80	50.3	4530	22	AA019731	Human tyrosine kin
34	80	50.3	4530	24	AB235012	Human gene express
35	80	50.3	4530	24	ABV94128	Breast carcinoma r
36	80	50.3	4530	24	ABN85585	Human HER2-neu SEQ
37	80	50.3	4530	24	ABK83918	Human cDNA differe
38	80	50.3	4530	25	ACC50139	Breast cancer asso
39	80	50.3	4530	25	ABQ83856	Human Her2/Neu enc
40	80	50.3	9274	22	AAF24297	HER2 transgene pla
41	80	50.3	9274	24	AA043934	HER-2 transgene pl
42	80	50.3	9274	24	ABK14057	Human HER2 (ErbB2)
43	79	49.7	1115	21	AA021778	Human breast and o
44	79	49.7	1755	24	AA032746	Human cDNA for the
45	79	49.7	1767	24	AA032744	Human cDNA for the

ALIGNMENTS

RESULT 1
ABA92252
ID ABA92252 standard; cDNA; 2763 BP.
XX
AC ABA92252;
XX
DT 17-JUN-2002 (first entry)
XX
DE Mouse Her-2/neu extracellular-phosphorylation domain fusion cDNA.
XX
KW Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
KW receptor; mouse; gene therapy; gene; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 1..2763


```

Qy      1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSerGlnAsnGluAspIleu 20
       ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1918 GAACGAGGCTGCACGAGCAGAGAGCCAGGCCCGACTGACGTCTCAGAACGAGGACTTA 1977
       |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy      21 GlyProAlaSerProLeuAspSerThrPhe 30
       |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1978 GGGCCCCCTCCAGCCCCCATGGACAGCACCTTC 2007
       |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 3
AAAT72725
ID      AAT72725 standard; cDNA; 2385 BP.
XX
AC      AAT72725;
XX
XX      17-SEP-1997 (first entry)
DE
XX      Her2-GM-CSF immunostimulant fusion protein DNA.
XX
KW      Her2-GM-CSF; granulocyte macrophage colony stimulating factor;
KW      growth factor receptor; oncogene; immunostimulant; cancer;
KW      therapy; ss.
XX
OS      Homo sapiens.
XX
FH      Key
CU      Location/Qualifiers
FT      11..2359
FT      /*tag= a
FT      /product= GM-CSF-Her2 fusion protein
FT      mRNA
FT      11..1969
FT      /*tag= b
FT      /product= Her2
FT      mRNA
FT      1970..1975
FT      /*tag= c
FT      /product= Leu-Glu linker
FT      mRNA
FT      1976..2359
FT      /*tag= d
FT      /product= GM-CSF
XX
XX      WO9724438-A1.
PN
XX
XX      10-JUL-1997.
XX
XX      23-DEC-1996; 96WO-US20241.
XX
XX      28-DEC-1995; 95US-0579823.
XX
XX      (ACTI-) ACTIVATED CELL THERAPY INC.
PA
XX      Laus R, Ruegg CL, Wu H;
PI
PI      WPI; 1997-363674/33.
DR      P-PSDB; AAU19764.
XX
XX      Potent APC that activates T-cells to give multivalent cellular
XX      immune response - can also induce a cytotoxic T-cell response in a
XX      vertebrate subject
XX
XX      Disclosure; Fig 8; 45pp; English.
XX
XX      A nucleic acid molecule (AAT72725) codes for a fusion protein
XX      (AAW19764), comprising granulocyte-macrophage colony stimulating
XX      factor (GM-CSF) and Her2, a growth factor receptor that is
XX      over-expressed in breast and ovarian cancer cells. It was
XX      prepd. by PCR amplification of Her2 cDNA from a breast cancer
XX      cell line and fusion to GM-CSF cDNA. Fusion expression vectors can
XX      be used to transfect mammalian and insect cells. The Her2-GM-CSF
XX      fusion protein is used to generate anti-Her2 immunity. Tumour
XX      cells are eliminated by cytotoxic T lymphocytes activated in vivo
XX      or in vitro by exposure to antigen-presenting cells exposed to the
XX      fusion protein.
XX
XX      Sequence 2385 BP; 488 A; 780 C; 677 G; 440 T; 0 other;
```


DE Humanised vector pITL-hHER/neu.
 XX Vector: vaccine; tumour; antigen; plasmid pITL-hHER/neu;
 KW human; HER-2/neu; C-erbB-2; breast cancer; ds.
 XX

OS Chimeric - Homo sapiens.
 OS Chimeric - Escherichia coli.
 OS Synthetic.

XX Key Location/Qualifiers
 FT CDS 13..921
 FT /tag= a
 FT /product= human HER-2/neu
 FT polyA_site 922..1181
 FT /tag= b
 FT /note= "combined splice and polyA sequences"
 FT CDS 1195..1401
 FT /tag= c
 FT /note= "SupF gene"
 FT misc_feature 1412..1864
 FT /tag= d
 FT /note= "ColE1 origin of replication"
 FT promoter 1877..2125
 FT /tag= e
 FT /note= "RANTES promoter"

XX WO9806863-A1.

XX 19-FEB-1998.

XX 14-AUG-1997; 97WO-US14305.

XX 14-AUG-1996; 96US-0023931.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Nelson EL, Nelson PJ;

XX WPI; 1998-159552/14.

XX Humanised polynucleotide vectors - comprising human derived promoter
 PT and sequence acceptance site, used for the production of vaccines
 XX
 XX Example 7; Page 41-42; 125pp; English.

XX Plasmid pITL-hHER2/neu comprises base vector pITL (see AAV21724) and
 CC a human HER-2/neu nucleic sequence. pITL-hHER2/neu was used to
 CC evaluate the toxicity of anti-tumour vaccination in rats, and in
 CC phase I and phase II trials to evaluate polynucleotide vaccination
 CC in advanced breast cancer. Novel humanised vectors, which can be
 CC based on pITL, comprise a human-derived promoter or mammalian
 CC homologue which is functional in mammalian target tissue and cells
 CC and a sequence acceptance site which accepts cDNA products from
 CC RT-PCR cloning. The vectors are non-replicating in mammalian cells
 CC but are capable of extended stable expression of the target
 CC sequence, generating an immune response in immunised individuals.
 CC The vectors selectively elicit immune responses to the target
 CC sequences with little or no immune response to the other components
 CC of the vectors.

XX Sequence 2125 BP; 449 A; 650 C; 598 G; 428 T; 0 other;

Alignment Scores:
 Pred. No.: 0.124 Length: 2125
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 19 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x AAV21727 (1-2125)

QY 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15

Db 112 GACAAGGCGTCCCGCGAGCAGAGAGCCAGCCCTCTGACGTC 156
 RESULT 7
 AAZ50586
 ID AAZ50586 standard; DNA; 2871 BP.
 XX
 AC AAZ50586;
 XX
 DT 23-MAY-2000 (first entry)
 XX
 DE DC8scFv-erbB2EC fusion construct containing tetramerisation domain.
 XX
 KW DC8 scFv; single-chain variable fragment; erbB2EC; extracellular domain;
 KW human; fusion construct; tetramerisation domain; constant domain;
 KW heteroinibitor; multifunctional compound; melanoma; sarcoma;
 KW immunoglobulin; cytostatic; immunostimulatory; antileukaemia; diagnosis;
 KW antiproliferative; prevention; treatment; malignant; haematopoietic cell;
 KW lymphoma; leukaemia; solid tumour; carcinoma; ds.
 XX
 OS Chimeric - Unidentified.
 OS Chimeric - Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 10..2865
 FT /tag= a
 FT /product= "DC8scFv-erbB2EC fusion protein"
 FT sig_peptide 10..66
 FT /tag= b
 FT /label= Leader_sequence
 FT misc_feature 67..390
 FT /tag= c
 FT /label= DC8scFv_light_chain_variable_region
 FT misc_feature 391..435
 FT /tag= d
 FT /label= Glycine-Serine-linker_DNA
 FT misc_feature 436..771
 FT /tag= e
 FT /label= DC8scFv_heavy_chain_variable_region
 FT misc_feature 775..807
 FT /tag= f
 FT /note= "5' end of human IgG3 upper hinge region
 FT with 3 additional nucleotides"
 FT misc_feature 808..924
 FT /tag= g
 FT /label= Human_p53_tetramerisation_domain
 FT misc_feature 925..945
 FT /tag= h
 FT /label= Short_peptide_linker
 FT misc_feature 946..2844
 FT /tag= i
 FT /label= erbB2EC_domain
 FT misc_feature 2845..2862
 FT /tag= j
 FT /label= His_tag
 XX WO200006605-A2.
 XX
 PD 10-FEB-2000.
 XX
 PF 28-JUL-1999; 99WO-EP05416.
 XX
 PR 28-JUL-1998; 98EP-0114082.
 XX
 PA (MITR-) MICROMET GES BIOMEDIZINISCHE FORSCHUNG.
 XX
 PI Kufer P, Dreier T, Baeuerle PA, Borschert K, Zettl F;
 XX WPI; 2000-195265/17.
 DR P-PSDB; AAY44993.
 DR
 XX New multifunctional compounds useful for preventing and/or treating
 PT malignant cell growth and for detection and diagnosis -

XX Example 9; Fig 49; 166pp; English.

PS The patent discloses heteroninibodies which are multifunctional compounds

CC producible in a mammalian host cell as a secretable and fully functional

CC heterodimer of two polypeptide chains, where one of the polypeptide

CC chains comprises, a CH1-domain (constant domain of an immunoglobulin

CC heavy chain) and the other chain comprises CH2-domain (constant domain of

CC an immunoglobulin light chain). The polypeptide chains further comprise,

CC fused to the constant domains at least two (polypeptide) chains having

CC different receptor or ligand functions, where further at least two of the

CC different (polypeptide) chains lack an intrinsic affinity for one another and

CC are linked via the constant domains. The heteroninibodies have

CC cytostatic, immunostimulatory, antileukaemia and antiproliferative

CC activities. These compounds can be used for diagnosing, preventing and

CC treating malignant cell growth related to malignancies of haematopoietic

CC cells e.g. lymphomas and leukaemias, or to solid tumours e.g. carcinomas,

CC melanomas and sarcomas.

CC The present sequence is a fusion construct comprising DC8

CC single-chain Fv (scFv) fragment at the N-terminus, extracellular

CC domain of human erbB2 at the C-terminus and a tetramerisation

CC domain between them. This construct was prepared to find out whether

CC an oligomerisation domain characterised in bacterial expression system

CC is applicable for expression of fully functional and secretable

CC recombinant protein in mammalian host cells. This tetrameric construct

CC was not expressed as secretable and fully functional protein

CC in mammalian cells. Hence general applicability of the tetramerisation

CC domain for oligomerisation strategies in mammalian cells was ruled out.

XX

SQ Sequence 2871 BP; 598 A; 868 C; 834 G; 571 T; 0 other;

Alignment Scores:

Pred. No.:	0.179	Length:	2871
Score:	80.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.31%	Indels:	0
DB:	21	Gaps:	0

JUNC_SEQ3_SEQ5 (1-30) x AAZ50586 (1-2871)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 2797 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTC 2841

RESULT 8

AAA89736

ID AAA89736 standard; DNA; 3600 BP.

XX

AC AAA89736;

XX

12-JAN-2001 (first entry)

DT

XX

Human HER-2/neu coding sequence.

DE

XX

Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;

KW breast cancer; prostate cancer; ovarian cancer; lung cancer;

KW colon cancer; ds.

KW

XX

OS Homo sapiens.

XX

Key

Location/Qualifiers

1..3600

FT CDS

FT /**tag- a

FT /product= "HER-2/neu protein"

XX

PN WO200044899-A1.

XX

03-AUG-2000.

XX

28-JAN-2000; 2000WO-US02164.

XX

29-JAN-1999; 99US-0117976.

XX

PR

XX (CORI-) CORIXA CORP.

PA (SMJK) SMITHKLINE BEECHAM.

XX

PI Cheever MA, Gheysen D;

XX

WPI; 2000-505976/45.

DR

P-PSDB; AAB21198, AAB21208.

XX

HER-2/neu extracellular domain/phosphorylation domain fusion proteins

PT useful for vaccinating against breast, ovarian, colon, lung and

PT prostate cancers -

XX

Disclosure; Fig 15; 128pp; English.

XX

The present sequence encodes the human HER-2/neu protein. HER-2/neu is a member of the tyrosine kinase family of receptor-like glycoproteins and shows homology to the epidermal growth factor receptor (EGFR). It probably plays a part in cell growth and/or differentiation. The HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation domain may be used to treat or prevent cancer by eliciting or enhancing an immune response to the HER-2/neu protein. It may be used to treat malignancies such as breast, ovarian, colon, lung and prostate cancers, and may be used as an antigen to vaccinate against these neoplasias.

SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 other;

Alignment Scores:

Pred. No.:	0.235	Length:	3600
Score:	80.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.31%	Indels:	0
DB:	21	Gaps:	0

JUNC_SEQ3_SEQ5 (1-30) x AAA89736 (1-3600)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTC 1959

RESULT 9

ABR86207

ID ABR86207 standard; cDNA; 3678 BP.

XX

AC ABR86207;

XX

24-SEP-2002 (first entry)

DT

XX

cDNA encoding human breast cancer antigen, Her2 variant.

DE

XX

Human; Her2; cytostatic; antiviral; immunostimulant;

KW cell-mediated immune response; tumour; breast cancer;

KW virus infection; prostate cancer; colorectal cancer; pancreatic cancer;

KW lymphoma; leukaemia; hepatitis virus; lentivirus; herpesvirus;

KW human immunodeficiency virus; HIV; flavivirus; pestivirus; gene; ss.

XX

OS Homo sapiens.

XX

Key

Location/Qualifiers

7..3678

FT CDS

FT /**tag- a

FT /product= "Breast cancer antigen Her2 variant"

XX

PN WO200240059-A2.

XX

23-MAY-2002.

XX

01-NOV-2001; 2001WO-US45626.

XX

01-NOV-2000; 2000US-0704232.

PR

XX (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.
 PA (MINC/) MINCHEFF M S.
 PA (LOUK/) LOUKINOV D I.
 PA (ZOUB/) ZOUBAK S.
 XX
 PI Mincheff MS, Loukinov DI, Zoubak S;
 XX
 DR WPI: 2002-527524/56.
 DR P-PSDB; AAU98923.
 XX
 XX Reducing a cell-mediated immune response against a target antigen,
 PT inducing undesired cells and stimulating presentation of an antigen by
 PT a cell, comprises administering a polynucleotide encoding a variant of
 PT an antigen.
 XX
 XX Disclosure: Page 128-134; 146pp; English.
 PS
 XX The invention relates to a method of inducing a cell-mediated immune
 CC response against a cell comprising a target antigen (I) in a subject,
 CC treating a subject having undesired cells, for example tumour cells
 CC or virally infected cells (C), reducing the number of (C) in a subject,
 CC and stimulating presentation of (I) by a cell. This is done by
 CC administering a polynucleotide (II) encoding a variant of (I), so that
 CC (II) expressed in a cell and cell-mediated immune response is induced.
 CC The method can be used to treat prostate cancer, breast cancer,
 CC colorectal cancer and pancreatic cancer, as well as lymphomas and
 CC leukaemias. The method is also useful in treating chronic viral
 CC infections such as those caused by hepatitisviruses, lentiviruses
 CC (including human immunodeficiency virus (HIV)), herpesviruses and the
 CC flaviviruses and pestiviruses. The present sequence represents the coding
 CC sequence of human breast cancer antigen, Her2 variant, used as a target
 CC antigen in the method of the invention.
 XX
 SQ Sequence 3678 BP; 746 A; 1137 C; 1091 G; 704 T; 0 other;

Alignment Scores:
 Pred. NO.: 0.241 Length: 3678
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 24 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x ABK86207 (1-3678)

OY 1 AspLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 Db 1825 GACAAGGGCTGCCCCCGCAGAGAGAGAGCCGCTCTGAGCTCC 1869

RESULT 10

AAT40739
 ID AAT40739 standard; cDNA; 3768 BP.

XX AAT40739;

XX 01-JAN-1997 (first entry)

XX HER-2/neu oncogene.

XX HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
 KW breast cancer; ovary cancer; colon cancer; lung cancer;
 KW prostate cancer; genetic immunisation; tumour; vaccine; vector;
 KW ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..3765

FT /*tag= b

FT /note= "nucleotides 2026-3765 (claim 1) code for
 FT HER-2/neu intracellular domain"

XX

PN W09630514-A1.
 XX
 PD 03-OCT-1996.
 XX
 EF 28-MAR-1996; 96WO-US01689.
 XX
 PR 31-MAR-1995; 95US-0414417.
 XX
 XX (UNIW) UNIV WASHINGTON.
 XX
 XX Cheever MA, Disis ML;
 PI WPI; 1996-455361/45.
 DR P-PSDB; AAW01111.
 DR
 XX DNA encoding HER-2/neu poly:peptide(s) - used for prevention or
 PT treatment of malignancies with which the HER-2/neu oncogene is
 PT associated
 PT
 PS Claim 1; Page 49-56; 71pp; English.
 XX
 XX Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or
 CC c-erbB2) protein (AAW01111). The oncogene is overexpressed in various
 CC cancers, including breast, ovarian, colon, lung and prostate, and
 CC appears to induce malignancies through quantitative mechanisms that
 CC result from increased or deregulated expression of an essentially
 CC normal gene product. Nucleotides 2026-3765 of the cDNA sequence
 CC code for the intracellular domain (Lys676-Val1255) of the HER-2/neu
 CC protein, which is useful for immunisation against malignancy.
 CC Nucleic acids can be used to direct expression of the intracellular
 CC domain in transformed host cells, or are used, alone or in a viral
 CC vector, for genetic immunisation of an animal.
 XX
 SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
 Pred. NO.: 0.248 Length: 3768
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 17 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x AAT40739 (1-3768)

OY 1 AspLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 Db 1915 GACAAGGGCTGCCCCCGCAGAGAGAGAGCCGCTCTGAGCTCC 1959

RESULT 11

AA01912
 ID AA01912 standard; DNA; 3768 BP.

XX AA01912;

XX 21-APR-1999 (first entry)

XX Human HER-2/neu oncogene DNA.

XX HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
 KW malignancy; treatment; tumour; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..3768

FT /*tag= a

FT /product= "HER-2/neu"

FT /note= "oncogene"

FT misc_feature 2026..3765

FT /*tag= b

FT /note= "region which elicits immune response"

XX

PN US5869445-A.
XX
PD
XX 09-FEB-1999.
XX
XX 01-APR-1996; 96US-0625101.
XX
PR 01-APR-1996; 96US-0625101.
PR 17-MAR-1993; 93US-0033644.
PR 12-AUG-1993; 93US-0106112.
PR 31-MAR-1995; 95US-0414417.
XX
XX (UNIW) UNIV WASHINGTON.
XX
XX Cheever MA, Disis ML;
PI WPI: 1999-152835/13.
DR P-PSDB; AAW92406.
XX
XX Use of HER-2/neu polypeptides - for eliciting an immune response to
PT an HER-2/neu associated malignancy, particularly for treating or
PT preventing tumours
XX
XX Claim 1a; Column 23-32; 26pp; English.
XX
XX This sequence encodes the human HER-2/neu oncogene protein. A fragment
CC of this protein is used in a method for eliciting or enhancing an immune
CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
CC B cells to produce an immune response to the HER-2/neu protein. The
CC method can be used for immunisation against a malignancy in which the
CC HER-2/neu oncogene is associated and in the treatment of an existing
CC tumour, or to prevent tumour occurrence or reoccurrence.
XX
XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 0.248 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 20 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x AAX01912 (1-3768)
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 12
AAA09455
ID AAA09455 standard; DNA; 3768 BP.
XX
XX AAA09455;
AC
XX 10-AUG-2000 (first entry)
XX
XX Human heregulin 2 (Her2) coding sequence.
XX
XX Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
KW self-protein; cell-associated peptide antigen; foreign epitope;
KW cancer; breast cancer; prostate cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
PR 20-OCT-1998; 98US-0105011.
XX

PA (MEBI-) M & E BIOTECH AS.
XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
DR WPI: 2000-349917/30.
DR P-PSDB; AAY92620.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
PS Claim 62; Page 187-193; 220pp; English.
XX
XX The claims detail a method for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, human
CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively.
XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 0.248 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x AAA09455 (1-3768)
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 13
AAH23392
ID AAH23392 standard; DNA; 3768 BP.
XX
XX AAH23392;
AC
XX 25-SEP-2001 (first entry)
DT
XX Human HER-2/neu protein encoding DNA.
DE
XX Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH 1..3768
FT CDS /*tag= a
FT /*product= "HER-2/neu protein"
XX
XX WO200153463-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US01850.
XX
XX 21-JAN-2000; 2000US-0177545.
XX

CC (ds) structure of not more than 49 consecutive nucleotides (nt), where
 CC at least a segment of one strand of the ds structure is complementary
 CC with the target gene and the cells are treated with interferon before
 CC introduction of dsRNA. The method is used to inhibit expression of
 CC target genes, particularly oncogenes, cytokine genes, id (not defined)
 CC protein genes; developmental or prion genes, or genes expressed in
 CC pathogenic organisms (particularly plasmodia) or in viruses or viroids
 CC (pathogenic in humans, animals or plants). Treating the cells with
 CC interferon greatly increases the extent to which dsRNA can inhibit
 CC expression of the target genes, and the effect is even greater when dsRNA
 CC are modified to increase their stability. ABX09936-ABX10075 represent
 CC gene fragments used to illustrate the method of the invention.

xx

SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:

Pred. No.:	0.248	Length:	3768
Score:	80.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.31%	Indels:	0
DB:	24	Gaps:	0

JUNC_SEQ3_SEQ5 (1-30) x ABX09987 (1-3768)

Qy	1	AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer	15
Db	1915	GACAGGGCTGCCCCCGCCGAGCAGAGAGCCCTCTGACGTCC	1959

Search completed: October 15, 2003, 23:48:41
 Job time : 31.4941 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 432.026 seconds
(without alignments)
2840.777 Million cell updates/sec

Title: JUNC_SEQ3_SEQ5

Perfect score: 159

Sequence: 1 DKGCPARQSRASPLTSONEDLGASPLDSTF 30

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/Cgn2_1/USPFO_SPOOL_P/HOLLERAN480/runat_15102003_131912_20536/app_query.fasta_1.4685
-DB=GenEmbl -QPMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=HOLLERAN480.CGN_1_1_22724@runat_15102003_131912_20536 -NCPU=6 -ICPU=3
-NO_MMWP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*

- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	144	90.6	2763	6	AX380942 Sequence
2	144	90.6	2781	6	AX380944 Sequence
C 3	85.5	53.8	186134	10	AL591390 Mouse DNA
C 4	85.5	53.8	214019	2	AC064803 Mus muscu
5	85	53.5	13450	9	AB096612 Homo sapi
6	85	53.5	16572	9	AB096614 Homo sapi
7	85	53.5	20271	9	AB096613 Homo sapi
8	85	53.5	30837	9	AY208911 Homo sapi
9	85	53.5	161815	9	AC079199 Homo sapi
10	85	53.5	168585	9	AC040933 Homo sapi
11	85	53.5	198008	2	AC142197 Homo sapi
12	84	52.8	2385	6	AR082744 Sequence
13	84	52.8	2385	6	AR099963 Sequence
14	84	52.8	2385	6	AR143949 Sequence
15	80	50.3	153	6	AX022984 Sequence
16	80	50.3	201	6	AR229723 Sequence
17	80	50.3	2871	6	AX023363 Sequence
18	80	50.3	3678	6	AX050514 Sequence
19	80	50.3	3768	6	AR034479 Sequence
20	80	50.3	3768	6	AX060704 Sequence
21	80	50.3	3768	6	AX201817 Sequence
22	80	50.3	3768	6	AX380923 Sequence
23	80	50.3	3768	6	AX384604 Sequence
24	80	50.3	3768	6	AX465456 Sequence
25	80	50.3	3768	6	AX467229 Sequence
26	80	50.3	3768	6	AX481438 Sequence
27	80	50.3	4473	6	AR080259 Sequence
28	80	50.3	4473	6	AR167390 Sequence
29	80	50.3	4473	9	HSEB2R
30	80	50.3	4530	6	AR202597 Sequence
31	80	50.3	4530	6	AR283481 Sequence
32	80	50.3	4530	6	AX282577 Sequence
33	80	50.3	4530	6	AX587649 Sequence
34	80	50.3	4530	6	AX644071 Sequence
35	80	50.3	4530	6	BD005474 Cellular
36	80	50.3	4530	6	I21124 Sequence 9
37	80	50.3	4530	6	I59745 Sequence 9
38	80	50.3	4530	9	HUMHER2A
39	80	50.3	9274	6	AX060703 Sequence
40	79	49.7	1755	6	AX384609 Sequence
41	79	49.7	1767	6	AX384607 Sequence
42	79	49.7	1773	6	AX384610 Sequence
43	79	49.7	1806	6	AX384608 Sequence
44	79	49.7	3780	4	AB008451 Canis fam
45	76	47.8	4062	10	D16295 Mesocricetu

ALIGNMENTS

RESULT 1

AX380942
LOCUS AX380942 2763 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 28 from Patent WO0212341.
ACCESSION AX380942
VERSION AX380942.1 GI:19575786
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS Cheever,M.A. and Gheysen,D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 28 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
Location/Qualifiers
source
1..2763
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD fusion protein cDNA"
BASE COUNT 571 a 855 c 772 g 565 t
ORIGIN
Alignment Scores:
Pred. No.: 8-97e-10 Length: 2763
Score: 144.00 Matches: 25
Percent Similarity: 100.00% Conservative: 5
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 90.57% Indels: 0
Gaps: 0
DB: 0
JUNC_SEQ3_SEQ5 (1-30) x AX380942 (1-2763)
QY 1 AsPLySGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1918 GAACGAGGTGCCACAGCAGACGAGCCAGCCAGTGCCTCTCAGAACGAGGACTTA 1977
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCTCCAGCCCATGGACAGCACCTTC 2007
RESULT 2
LOCUS AX380944 2781 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 30 from Patent WO0212341.
ACCESSION AX380944
VERSION AX380944.1 GI:19575787
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS Cheever,M.A. and Gheysen,D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 30 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
Location/Qualifiers
source
1..2781
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD-Tcp0 fusion protein cDNA"
BASE COUNT 574 a 859 c 779 g 569 t
ORIGIN
Alignment Scores:
Pred. No.: 9.03e-10 Length: 2781
Score: 144.00 Matches: 25
Percent Similarity: 100.00% Conservative: 5
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 90.57% Indels: 0
Gaps: 0
DB: 0

JUNC_SEQ3_SEQ5 (1-30) x AX380944 (1-2781)
QY 1 AsPLySGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1918 GAACGAGGTGCCACAGCAGACGAGCCAGCCAGTGCCTCTCAGAACGAGGACTTA 1977
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCTCCAGCCCATGGACAGCACCTTC 2007
RESULT 3
LOCUS AL591390 186134 bp DNA linear ROD 17-NOV-2001
DEFINITION Mouse DNA sequence from clone RP23-355L10 on chromosome 11,
complete sequence.
ACCESSION AL591390
VERSION AL591390.8 GI:17017767
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 186134)
AUTHORS Phillimore,B.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Nov 20, 2001 this sequence version replaced gi:16555512.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-355L10 is
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
RP23-355L10. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP23-355L10 is at 186134 in this
sequence. The true right end of clone RP23-438D7 is at 2000 in this
sequence.
FEATURES
Location/Qualifiers
source
1..186134
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="11"
/clone="RP23-355L10"
/clone_lib="RPCI-23"
17991..18044
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."
47794..47812
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."
BASE COUNT 46287 a 47900 c 47689 g 44258 t
ORIGIN

Alignment Scores:

```

Pred. No.:      5.11      Length:      186134
Score:          85.50      Matches:      18
Percent Similarity: 72.41%      Conservative: 3
Best Local Similarity: 62.07%      Mismatches: 5
Query Match:      53.77%      Indels:      3
DB:              10      Gaps:      1

```

JUNC_SEQ3_SEQ5 (1-30) x AL591390 (1-186134)

```

Qy 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21
    :::::|||||
Db 70080 GAGGCTGT-----CACCTCGTGTCTCCCTTCTTCTAGACGAGGACTTAGGC 70030

Qy 22 ProAlaSerProLeuAspSerThrPhe 30
    ||:::|||||:|||||:|||||
Db 70029 CCCTCCAGCCCCATGGACAGCACCTTC 70003

RESULT 4
AC064803/c
LOCUS      AC064803      214019 bp      DNA      linear      HTG 02-NOV-2001
DEFINITION Mus musculus chromosome 11 clone RP23-16G14 map 11, WORKING DRAFT
SEQUENCE, 14 unordered pieces.
AC064803
VERSION    AC064803.4      GI:16596607
KEYWORDS   HTG; HTGS, PHASE1; HTGS, DRAFT; HTGS, FULLTOP.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus

REFERENCE
1 (bases 1 to 214019)
AUTHORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boquslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collamore,A., Cooke,P., DeRellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 2, 2001 this sequence version replaced gi:12061538.
All repeats were identified using RepeatMasker:
Snit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9173
Center clone name: 16_G.14
----- Summary Statistics

```

Sequencing vector: M13; M77815; 54% of reads
 Sequencing vector: Plasmid; n/a; 46% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 208571 bases at least Q40
 Consensus quality: 210746 bases at least Q30
 Consensus quality: 211706 bases at least Q20
 Insert size: 197000; agarose-fp
 Insert size: 212719; sum-of-contigs
 Quality coverage: 8.5 in Q20 bases; agarose-fp
 Quality coverage: 7.8 in Q20 ba.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1 21450: contig of 21460 bp in length
* 21461 21560: gap of 100 bp
* 21561 21845: contig of 3285 bp in length
* 21846 24945: gap of 100 bp
* 24946 28115: contig of 3170 bp in length
* 28116 28215: gap of 100 bp
* 28216 35345: contig of 7130 bp in length
* 35346 35445: gap of 100 bp
* 35446 43994: contig of 8549 bp in length
* 43995 44094: gap of 100 bp
* 44095 83787: contig of 39693 bp in length
* 83788 83887: gap of 100 bp
* 83888 96462: contig of 12575 bp in length
* 96463 96562: gap of 100 bp
* 96563 108784: contig of 12222 bp in length
* 108785 108884: gap of 100 bp
* 108885 122875: contig of 13991 bp in length
* 122876 122975: gap of 100 bp
* 122976 139026: contig of 16051 bp in length
* 139027 139126: gap of 100 bp
* 139127 159729: contig of 20603 bp in length
* 159730 159829: gap of 100 bp
* 159830 178173: contig of 18344 bp in length
* 178174 178273: gap of 100 bp
* 178274 200970: contig of 22697 bp in length
* 200971 201070: gap of 100 bp
* 201071 214019: contig of 12949 bp in length.

```

FEATURES

```

source
1..214019
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/map="11"
/clone="RP23-16G14"
/clone_lib="RPCI-23 Female Mouse BAC"
misc_feature
1..21460
/note="assembly_fragment
clone_end:SP6
vector_side:left"
misc_feature
21561..24845
/note="assembly_fragment"
misc_feature
24946..28115
/note="assembly_fragment"
misc_feature
28216..35345
/note="assembly_fragment"
misc_feature
35446..43994
/note="assembly_fragment"
misc_feature
44095..83787
/note="assembly_fragment"
misc_feature
83888..96462
/note="assembly_fragment"
misc_feature
96563..108784
/note="assembly_fragment"

```

```

misc_feature 10885..122875
/note="assembly_fragment"
misc_feature 122976..139026
/note="assembly_fragment"
misc_feature 139127..159729
/note="assembly_fragment"
misc_feature 159830..178173
/note="assembly_fragment"
misc_feature 178274..200970
/note="assembly_fragment"
misc_feature 201071..214019
/note="assembly_fragment"
clone_end:T7
vector_side:right
BASE COUNT 52809 a 53348 c 53260 g 53244 t 1358 others
ORIGIN

Alignment Scores:
Pred. No.: 5.89 Length: 214019
Score: 85.50 Matches: 18
Percent Similarity: 72.41% Conservatives: 3
Best Local Similarity: 62.07% Mismatches: 5
Query Match: 53.77% Indels: 3
DB: 2 Gaps: 1

JUNC_SEQ3_SEQ5 (1-30) x AC064803 (1-214019)
QY 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21
Db 53608 GAGGGCTGT-----CACCTCGTGTCTCCCTTCCTTAGACGAGGACTTAGGC 53558

QY 22 ProAlaSerProLeuAspSerThrPhe 30
Db 53557 CCTCCAGCCCATGACAGACCTTC 53531

RESULT 5
AB096612
LOCUS AB096612 13450 bp DNA linear PRI 21-MAY-2003
DEFINITION Homo sapiens DNA, 13kb-normal EcoRI sequence.
ACCESSION AB096612
VERSION AB096612.1 GI:30962525
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Kuwahara,Y., Tanabe,C., Ikeuchi,T., Aoyagi,K., Nishigaki,M.,
Sakamoto,H., Hoshinaga,K., Yoshida,T., Sasaki,H. and Terada,M.
TITLE Presence of novel mechanisms other than the BFB cycles in
amplification of human oncogene
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 13450)
Sasaki,H.
DIRECT SUBMISSION
SUBMITTED (20-NOV-2002) Hiroki Sasaki, National Cancer Center
Research Institute, Genetics Division; Tsukiji 5-1-1, Chuo-ku,
Tokyo 104-0045, Japan (E-mail:hksasaki@ncc.ncc.go.jp,
Tel:81-3-3542-2511(ex.4402), Fax:81-3-3541-2685)
FEATURES
source
1..13450
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q12"
/clone="13kb-normal EcoRI DNA"
684..2046
/gene="c-ERBB-2"
684..936
/gene="c-ERBB-2"
684..936
/gene="c-ERBB-2"
1078..2046
/gene="c-ERBB-2"
/gene="c-ERBB-2"

misc_feature 10885..122875
/note="assembly_fragment"
misc_feature 122976..139026
/note="assembly_fragment"
misc_feature 139127..159729
/note="assembly_fragment"
misc_feature 159830..178173
/note="assembly_fragment"
misc_feature 178274..200970
/note="assembly_fragment"
misc_feature 201071..214019
/note="assembly_fragment"
clone_end:T7
vector_side:right
BASE COUNT 52809 a 53348 c 53260 g 53244 t 1358 others
ORIGIN

Alignment Scores:
Pred. No.: 5.89 Length: 214019
Score: 85.50 Matches: 18
Percent Similarity: 72.41% Conservatives: 3
Best Local Similarity: 62.07% Mismatches: 5
Query Match: 53.77% Indels: 3
DB: 2 Gaps: 1

JUNC_SEQ3_SEQ5 (1-30) x AC064803 (1-214019)
QY 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21
Db 53608 GAGGGCTGT-----CACCTCGTGTCTCCCTTCCTTAGACGAGGACTTAGGC 53558

QY 22 ProAlaSerProLeuAspSerThrPhe 30
Db 53557 CCTCCAGCCCATGACAGACCTTC 53531

RESULT 5
AB096612
LOCUS AB096612 13450 bp DNA linear PRI 21-MAY-2003
DEFINITION Homo sapiens DNA, 13kb-normal EcoRI sequence.
ACCESSION AB096612
VERSION AB096612.1 GI:30962525
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Kuwahara,Y., Tanabe,C., Ikeuchi,T., Aoyagi,K., Nishigaki,M.,
Sakamoto,H., Hoshinaga,K., Yoshida,T., Sasaki,H. and Terada,M.
TITLE Presence of novel mechanisms other than the BFB cycles in
amplification of human oncogene
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 13450)
Sasaki,H.
DIRECT SUBMISSION
SUBMITTED (20-NOV-2002) Hiroki Sasaki, National Cancer Center
Research Institute, Genetics Division; Tsukiji 5-1-1, Chuo-ku,
Tokyo 104-0045, Japan (E-mail:hksasaki@ncc.ncc.go.jp,
Tel:81-3-3542-2511(ex.4402), Fax:81-3-3541-2685)
FEATURES
source
1..13450
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q12"
/clone="13kb-normal EcoRI DNA"
684..2046
/gene="c-ERBB-2"
684..936
/gene="c-ERBB-2"
1078..2046
/gene="c-ERBB-2"
/gene="c-ERBB-2"

misc_feature 10885..122875
/note="3' exon"
11349..11517
/gene="GRB7"
11349..11517
/gene="GRB7"
11349..11517
/gene="GRB7"
BASE COUNT 2748 a 3818 c 3387 g 3496 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 0.415 Length: 13450
Score: 85.00 Matches: 23
Percent Similarity: 46.00% Conservatives: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match: 53.46% Indels: 24
DB: 9 Gaps: 2

JUNC_SEQ3_SEQ5 (1-30) x AB096612 (1-13450)
QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
Db 96 CCTGCAGAGGGTGGGAGGAGAGATGAGTCCAGATATGCCAGGCCCTCAGGAGGCTGC 155
QY 16 -----GlnAsnGluAspLeu 20
Db 156 ATGCTGGGTGGGGAGGGGCCACCATCTCTCTCTCTCCACAGATGAGGACTTG 215
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 216 GGGCCAGCAGTCCCTTGGACAGCACCTTC 245

RESULT 6
AB096614
LOCUS AB096614 16572 bp DNA linear PRI 21-MAY-2003
DEFINITION Homo sapiens DNA, 16.5kb-T5 rearranged sequence.
ACCESSION AB096614
VERSION AB096614.1 GI:30962527
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Kuwahara,Y., Tanabe,C., Ikeuchi,T., Aoyagi,K., Nishigaki,M.,
Sakamoto,H., Hoshinaga,K., Yoshida,T., Sasaki,H. and Terada,M.
TITLE Presence of novel mechanisms other than the BFB cycles in
amplification of human oncogene
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 16572)
Sasaki,H.
DIRECT SUBMISSION
SUBMITTED (20-NOV-2002) Hiroki Sasaki, National Cancer Center
Research Institute, Genetics Division; Tsukiji 5-1-1, Chuo-ku,
Tokyo 104-0045, Japan (E-mail:hksasaki@ncc.ncc.go.jp,
Tel:81-3-3542-2511(ex.4402), Fax:81-3-3541-2685)
FEATURES
source
1..16572
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q12"
/clone="16.5kb-T5 rearranged DNA"
684..2046
/gene="c-ERBB-2"
684..936
/gene="c-ERBB-2"
1078..2046
/gene="c-ERBB-2"
11349..11517
/gene="GRB7"
11349..11517
/gene="GRB7"

```

[illegible]

variation	902	/frequency="0.01"				/rpt_type=dispersed
variation	975	/replace="a"				4784. .4858
gene		/frequency="0.03"				/rpt_family="MIR"
mRNA		/replace="t"				/rpt_type=dispersed
		/genes="ERBB2"				5404. .5693
		join(1253. .1421,8102. .8253,9440. .9653,10437. .10571,				/rpt_family="Alu"
		10932. .11000,11205. .11320,11459. .11600,13047. .13166,				/rpt_type=dispersed
		13441. .13567,16405. .16478,16585. .16855,16859. .17058,				6074. .6282
		17420. .17552,17634. .17724,18439. .18599,20908. .20955,				/rpt_family="MIR"
		24439. .24577,24658. .24780,25032. .25130,25858. .26043,				/rpt_type=dispersed
		26181. .26336,26459. .26534,26839. .26985,27694. .27791,				6952. .6953
		27947. .28135,28427. .28679,28821. .29789)				/genes="ERBB2"
		/product="v-erb-b2 erythroblastic leukemia viral oncogene				/replace="t"
		homolog 2, neuro/glioblastoma derived oncogene homolog				/frequency="0.98"
		(avian)"				6957
CDS		join(1349. .1421,8102. .8253,9440. .9653,10437. .10571,				/genes="ERBB2"
		10932. .11000,11205. .11320,11459. .11600,13047. .13166,				/frequency="0.02"
		13441. .13567,16405. .16478,16585. .16855,16859. .17058,				/replace="g"
		17420. .17552,17634. .17724,18439. .18599,20908. .20955,				6960
		24439. .24577,24658. .24780,25032. .25130,25858. .26043,				/genes="ERBB2"
		26181. .26336,26459. .26534,26839. .26985,27694. .27791,				/frequency="0.02"
		27947. .28135,28427. .28679,28821. .29176)				/replace="g"
		/genes="ERBB2"				7036
		/codon_start=1				/genes="ERBB2"
		/product="v-erb-b2 erythroblastic leukemia viral oncogene				/frequency="0.03"
		homolog 2, neuro/glioblastoma derived oncogene homolog				/replace="t"
		(avian)"				7146
		/protein_id="AA018082.1"				/genes="ERBB2"
		/db_xref="GI:27658000"				/frequency="0.01"
		/translation="MELAAACRWGLLLALLPPCAASTOVCTGCTDMKRLPASPETHLD				/replace="t"
		MLRHLYQGVQGNLELYLPTNASLSFLQIQEVOGYVLIAHNOVRVPIQLRLIV				7151
		RGTLDFNLALVDNDGDFLNNTPVTGASPGGURELQRLSLTEILKGGVLIQRNPQ				/genes="ERBB2"
		LCYQDILFMKNQALALTIDITNSTRACHPKSCMKRCWGSESSDCQSLRT				/frequency="0.01"
		VCAGGACRCPGLPTDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNT				/replace="t"
		DTFESMPNPGRTVFCAGCVTACPVNLTSDVSCVLCPLHNOEVTADGQIQCEKC				7825
		SKPCARVCYGLGMEHLREYRVTSAIOEFAGCKIEGSLAFLPESFDGDPASNAPL				/genes="ERBB2"
		QPEQLQFTLEITGILYISAMPDLSPLVFQNLQVIRGRILLNGAYSLTLQGLGI				/frequency="0.06"
		SWGLSLRELGSGLAIHHNTHLCFVHTVPMDFRNPHQALLHTANRPEDECVEG				/replace="t"
		LACHQACVGHGPGTQCVCNCSQFLRQCEVCRCVQLQGLPREYVNAHRLCPHE				7874
		COFQNGSVCFGEADQCACAHYKDPFCVACRCPGDLSPYIMKFPDEEACQ				/genes="ERBB2"
		PCPINTHSCVDLDKDCPAEQRASPLTSIIISAVGILLVVLVGVFGILIKRROKI				/frequency="0.01"
		RKYTMRLQLQETLEPLTPSGAMPNQAOMRIILKETLRKVKVLGSGARGTVYKGIWI				/replace="a"
		PDGENVKIPVALIKVLENTSPKANKELIDEAIVMAGVSPYVSRLLIGLTLTVQLVT				8127
		QLMPYCLLDHVRNRRGSLGQDLNWCMIAGMSILEDVRLVHRDLAARNVLYKSP				/genes="ERBB2"
		NHVKITDFGLALDIDETEYHADGKVPYIKWMALESILRRRFTHQSDVMSYGVYWE				/frequency="0.01"
		LMTFGAKPYDIPAREIPDLLEKGERLPQPICTIDVYIMVYKWMIDSECRPREL				/replace="a"
		VSEFSMRARDPQRFVYVIONEDLGPASPLDTFVRSLLDDMDGLVDAEYLVPOQGF				8467. .8473
		FCPDAPAGAGVHHRHSRSTRSGGDLTLGLEPSEEAAPRSLAPSGAGSDVFG				/genes="ERBB2"
		DIGMNAKGLQSLPHDPSPLARYSEDVPLPSETDGVAPLTSPQPEYVNPQDVR				/frequency="0.96"
		POPPSPREGPLAARPAAGATLERAKTLSPGKNGVVKDFAFGGAVENPEYILTPQGGA				/replace="t"
		POPPPPPAFSPAFDNLNYYWDQPPPERGAPPSTFKGTPTAENPEYILGLDVPV"				8574
		1422. .1712				/genes="ERBB2"
misc_feature		/genes="ERBB2"				/frequency="0.02"
		/note="Region not scanned for variation"				/replace="t"
variation		1839				8575. .8707
		/genes="ERBB2"				/rpt_family="MIR"
		/frequency="0.02"				/rpt_type=dispersed
		/replace="c"				8836. .9130
misc_feature		1887. .6576				/rpt_family="Alu"
		/genes="ERBB2"				/rpt_type=dispersed
		/note="Region not scanned for variation"				8875. .9119
repeat_region		2315. .2536				/genes="ERBB2"
		/rpt_family="MIR"				/note="Region not scanned for variation"
repeat_region		/rpt_type=dispersed				9136. .9214
		3217. .3417				/rpt_family="MIR"
repeat_region		/rpt_family="Alu"				/rpt_type=dispersed
		/rpt_type=dispersed				9595
		3837. .3974				/genes="ERBB2"
		/rpt_family="MIR"				/frequency="0.02"

repeat_region	/replace="g" 9702..9770 /rpt_family="Tip100" /rpt_type=dispersed 9766
variation	/gene="ERBB2" /frequency="0.01" /replace="t" 9799..9946
repeat_region	/rpt_family="Tip100" /rpt_type=dispersed 9949..10151
repeat_region	/rpt_family="MIR" /rpt_type=dispersed 10009
variation	/gene="ERBB2"

Alignment Scores:		
Pred. No.:	0.963	30937
Score:	85.00	23
Percent Similarity:	46.00%	0
Best Local Similarity:	46.00%	3
Query Match:	53.46%	24
DB:	9	2
		Length:
		Matches:
		Conservative:
		Mismatches:
		Indels:
		Gaps:

JUNC_SEQ3_SEQ5 (1-30) x AY208911 (1-30837)

QY	5	ProAlaGlu-----GlnArgAlaSerProLeuThrSer-----	15
Db	27839	CCTGCAGAGGGTGGGAAGGAGAGATGAGTCCAGTATGCCAGGCCCTCCGGAAGCTGC	27898
QY	16	-----GlnAsnGluAspLeu	20
Db	27899	ATGCTGGGTGGGGGGGCCACCATCTGCTCTCTTCTCCACAGAAATGAGGACTTG	27958
QY	21	GlyProAlaSerProLeuAspSerThrPhe	30
Db	27959	GGCCCAAGCAGTCCCTTGGACAGCACTTC	27988

RESULT	9
AC079199	
LOCUS	161815 bp DNA linear
DEFINITION	Homo sapiens chromosome 17, clone RP11-94L15, complete sequence.
ACCESSION	AC079199
VERSION	AC079199.9 GI:23462913
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 161815)

AUTHORS Birren, B., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 17, clone RP11-94L15
 REFERENCE 2 (bases 1 to 161815)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E.

Anderson, S., Barna, N., Bestien, V., Beda, F., Boguslavsky, L., Boukhalper, B., Brown, A., Burkett, G., Campolano, A., Castle, A., Choepel, Y., Collangelo, M., Collins, S., Collamore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Ferrelira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRocque, K., Lamarez, R., Landers, T., Lenockzy, J., Levine, R., Lieu, C., Liu, G., MacDonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meidrum, J., Meneus, L., Mihova, T., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sougnez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Triglio, J., Vassiliev, H., Viel, R., Vo, A.,

TITLE	JOURNAL	REFERENCE	AUTHORS
-------	---------	-----------	---------

Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (23-AUG-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 161815)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choeplet, J., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuppback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tsefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Willson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (08-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 161815)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, V., Collymore, A., Cook, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J., Sore, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Geir, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Ton, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Milenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Reita, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (02-OCN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 2, 2002 this sequence version replaced gi:22759800.
All repeats were identified using RepeatMasker:
Smit, A F A & Green, P (1996-1997).

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10635

FEATURES source

```
Location/Qualifiers
1. .161815
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="17"
   /map="17"
   /clone="RP11-94L15"
   /clone_lib="RPC1-11 Human"
7. .470
   /note="PCR product sequence"
```

unsure

```

repeat_region complement(262..317)
unsure /rpt_family="MIR"
repeat_region /note="PCR product sequence only"
repeat_region 1689..1986
repeat_region 5033..5179
repeat_region 6113..6233
repeat_region /rpt_family="MIR"
repeat_region 6418..6460
repeat_region /rpt_family="L3"
repeat_region 7910..7930
repeat_region /rpt_family="AT-rich"
repeat_region 8422..8476
repeat_region /rpt_family="(CCTG)n"
repeat_region 8513..8659
repeat_region /rpt_family="MIR"
repeat_region complement(8917..9066)
repeat_region /rpt_family="MIR3"
repeat_region 12588..12697
repeat_region /rpt_family="G-rich"
repeat_region complement(13513..13824)
repeat_region /rpt_family="AluSp"
repeat_region complement(13836..13937)
repeat_region 18739..18939
repeat_region /rpt_family="MIR"
repeat_region 21075..21274
repeat_region /rpt_family="MIR"
repeat_region 21447..21714
repeat_region /rpt_family="MIR"
repeat_region complement(23385..23684)
repeat_region /rpt_family="AluX"
repeat_region 24620..24819
repeat_region /rpt_family="MIR"
repeat_region 24991..25298
repeat_region /rpt_family="L2"
repeat_region complement(25299..25605)
repeat_region /rpt_family="AluJb"
repeat_region 25606..25686
repeat_region /rpt_family="L2"
repeat_region 25687..25838
repeat_region /rpt_family="MIR"
repeat_region 25839..25952
repeat_region /rpt_family="LTR54B"
repeat_region complement(25953..26241)
repeat_region /rpt_family="AluSp"
repeat_region 26242..26378
repeat_region /rpt_family="LTR54B"
repeat_region 26379..26695
repeat_region /rpt_family="AluY"
repeat_region 26696..26901
repeat_region /rpt_family="LTR54B"
repeat_region 26902..26985
repeat_region /rpt_family="MIR"
repeat_region 26986..27224
repeat_region /rpt_family="L2"
repeat_region 27243..27550
repeat_region /rpt_family="AluSg"
repeat_region 27695..28007
repeat_region /rpt_family="L2"
repeat_region 28290..28600
repeat_region /rpt_family="AluX"
repeat_region 29462..29600
repeat_region /rpt_family="MIR"
repeat_region 29643..29967
repeat_region /rpt_family="AluY"
repeat_region complement(30700..31000)
repeat_region /rpt_family="AluSp"
repeat_region complement(31002..31248)
repeat_region /rpt_family="AluJb"
repeat_region 31521..31572

```

```

/rpt_family="L2"
31577..31796
/rpt_family="MIR"
32751..33072
/rpt_family="L2"
33181..33293
/rpt_family="L2"
34072..34376
/rpt_family="AluSp"
complement(34542..34833)
/rpt_family="AluJb"
complement(35218..35297)
/rpt_family="L2"
35320..35431
/rpt_family="AluSp"
35433..35500
/rpt_family="AluSp/q"
complement(35581..35702)
/rpt_family="L2"
35710..36008
/rpt_family="AluSc"
36009..36023

Alignment Scores:
Pred. No.: 5.18 Length: 161815
Score: 85.00 Matches: 23
Percent Similarity: 46.00% Conservative: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match: 53.46% Indels: 24
DB: 9 Gaps: 2

JUNC_SEQ3_SEQ5 (1-30) x AC079199 (1-161815)
QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
Db 71166 CCTGCAGAGGGTGGGAAGGAGAGATGAGTCCAGCGCCCTCACGGAGGCTGC 71225
QY 16 -----GlnAsnGluAspLeu 20
Db 71226 ATGCTGGGCTGGGAGGGCCACCATCTGCTCTCTCTCCACAGATGAGGACTTG 71285

RESULT 10
AC040933 AC040933 168585 bp DNA linear PRI 01-OCT-2002
LOCUS Homo sapiens chromosome 17, clone CTD-2019C10, complete sequence.
DEFINITION AC040933
ACCESSION AC040933
VERSION AC040933.12 GI:23396303
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 168585)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone CTD-2019C10
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 168585)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bedalov,F.,
Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castelle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,

```

Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE JOURNAL

Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 168585)

REFERENCE AUTHORS

Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (01-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 168585)

TITLE JOURNAL

REFERENCE AUTHORS

Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (01-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 1, 2002 this sequence version replaced gi:22597589.

COMMENT

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8894
Center clone name: 2019_C_10

Only the last 167.6 kb of this clone are being submitted.
The remainder overlaps accession number AC097491 [WICGR project L11967].

FEATURES	Location/Qualifiers
source	1. 168585 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="17" /map="17"
repeat_region	/clone="CTD-2019C10" /complement(1359..1516) /rpt_family="MIR"
repeat_region	/complement(1538..1754) /rpt_family="MIR"
repeat_region	2117..2419 /rpt_family="AluSq" /complement(2609..2748) /rpt_family="MIR"
repeat_region	4091..4278 /rpt_family="MIR"
unsure	/note="<30 qual SINGL region" /complement(4334..4339)
repeat_region	4409..4725 /rpt_family="AluJb"
repeat_region	4782..4942 /rpt_family="L1ME1"
repeat_region	4986..5046 /rpt_family="(TGAA)n"
repeat_region	/complement(5492..5657) /rpt_family="MIR"
repeat_region	6438..6626 /rpt_family="CT-rich"
repeat_region	8771..8827 /rpt_family="CA-rich"
repeat_region	10677..10875 /rpt_family="MIR3"
repeat_region	/complement(12314..12419) /rpt_family="L2"
repeat_region	12443..12588 /rpt_family="AluJo"
repeat_region	12589..12685 /rpt_family="AluSp"
repeat_region	12886..13057 /rpt_family="AluJo"
unsure	/note="<30 qual SINGL region" /complement(14521..14553)
unsure	/note="<30 qual SINGL region" /complement(14571..14592)
unsure	/note="<30 qual SINGL region" /complement(14609..14614)
unsure	/note="<30 qual SINGL region" /complement(14624..14625)
unsure	/note="<30 qual SINGL region" /complement(14639..14666) /note="<30 qual SINGL region" 15796..15849
repeat_region	/rpt_family="GA-rich"
repeat_region	16715..17025 /rpt_family="AluSx"
repeat_region	/complement(17213..17283) /rpt_family="MIR"
unsure	/note="<30 qual SINGL region" /complement(17647..17651)
repeat_region	19655..20204 /rpt_family="L2"
repeat_region	20443..20649 /rpt_family="MIR"
repeat_region	20663..20928 /rpt_family="AluJo"
unsure	21295..21301 /note="<30 qual SINGL region" /complement(21929..21987)
repeat_region	/rpt_family="L2"
repeat_region	/complement(22041..22147) /rpt_family="MIR"

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 |||||
 Db 10 GACAGGGCTGCCCGCCGAGCAGAGAGCCCTCTGACGTCC 54

Search completed: October 16, 2003, 11:01:09
 Job time : 507.208 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 234.794 Seconds
(without alignments)
3105.423 Million cell updates/sec

Title: JUNC_SEQ3_SEQ4
Perfect score: 159
Sequence: 1 DRGCPAEQASPLTSQDNEDGASPLDSTF 30

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=framet_p2n.model -DEV=xlp
-Q=/cgp2_1/USPTO.spool_p/HOLLERAN480/runat_15102003_131913_20548/app_query.fasta_1.4685
-DB=EST -QPMI=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480@cgn_1.1.12645 @runat_15102003_131913_20548 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_esti:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	85.5	53.8	618	28	AZ699313	AZ699313 RPCI-23-2
2	85	53.5	422	9	AW606421	AW606421 RC3-HT037
3	85	53.5	439	10	BE157169	BE157169 RC3-HT037
4	85	53.5	944	14	CD558559	CD558559 AGENCOURT
5	81	50.9	458	10	BE157103	BE157103 RC3-HT037
6	80	50.3	378	10	BF869854	BF869854 IL3-ET011
7	80	50.3	411	2	HS069572	Bx479682 Homo sapi
8	80	50.3	433	10	BF762317	BF762317 IL2-CS004
9	80	50.3	450	12	BG988031	BG988031 MR2-HT116
10	80	50.3	466	12	BM720186	BM720186 UT-E-E00-
11	80	50.3	546	12	BI194790	BI194790 602948150
12	80	50.3	547	9	AA443351	AA443351 zw85b05.r
13	80	50.3	561	12	BM720098	BM720098 UT-E-E00-
14	80	50.3	607	12	BM083444	BM083444 imageqc_6
15	80	50.3	613	9	AW410534	AW410534 fh06h06.x
16	80	50.3	666	13	BX113765	BX113765 BX113765
17	80	50.3	667	12	BM048624	BM048624 603623984
18	80	50.3	674	12	BM758528	BM758528 K-EST0038
19	80	50.3	741	10	BE615590	BE615590 601278973
20	80	50.3	881	9	AI906106	AI906106 RC-BT105-
21	80	50.3	886	13	BQ882315	BQ882315 AGENCOURT
22	80	50.3	893	14	CA487631	CA487631 AGENCOURT
23	80	50.3	905	13	BU541491	BU541491 AGENCOURT
24	80	50.3	909	10	BG289457	BG289457 602381495
25	80	50.3	922	14	CA453990	CA453990 AGENCOURT
26	80	50.3	1009	13	BQ647343	BQ647343 AGENCOURT
27	79	49.7	278	10	BF923905	BF923905 QV4-NT025
28	79	49.7	294	10	BF757857	BF757857 QV4-CT057
29	79	49.7	337	10	BE091682	BE091682 IL2-BT073
30	79	49.7	348	9	AI909847	AI909847 QV-BT225-
31	79	49.7	378	14	CB266492	CB266492 1005398 H
32	79	49.7	423	12	BM703962	BM703962 UT-E-CK1-
33	79	49.7	464	2	HS067221	Bx477782 Homo sapi
34	79	49.7	496	12	BM795191	BM795191 K-EST0076
35	79	49.7	505	12	BM854421	BM854421 K-EST0136
36	79	49.7	531	12	BM830170	BM830170 K-EST0103
37	79	49.7	536	14	CB129403	CB129403 K-EST0179
38	79	49.7	547	12	BM787824	BM787824 K-EST0066
39	79	49.7	567	2	HS077766	Bx487579 Homo sapi
40	79	49.7	574	12	BM829991	BM829991 K-EST0103
41	79	49.7	630	2	HS073298	Bx483130 Homo sapi
42	79	49.7	691	14	CB853376	CB853376 UT-CF-FNO
43	79	49.7	767	10	BF240297	BF240297 601905830
44	79	49.7	849	14	CA489799	CA489799 AGENCOURT
45	79	49.7	852	14	CD516283	CD516283 AGENCOURT

ALIGNMENTS

RESULT 1
AZ699313
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

AZ699313
RPCI-23-240N16.TJ RPCI-23 Mus musculus genomic clone
, genomic survey sequence.
AZ699313
GI:12419623
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 618)

618 bp
DNA
linear
GSS 24-JAN-2001

AUTHORS	Zhao, S., Niernman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., Megann, S., Tsegaye, G., Geer, K., Kroi, M., de Jong, P. and Fraser, C.M.
TITLE	Mouse BAC End Sequences from Library RPCI-23
JOURNAL	Unpublished
COMMENT	Other GSs: RPCI-23-240N16.TV Contact: Shaying Zhao Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: szhao@tigr.org Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). clones may be purchased from BACPAC resources (http://www.choiri.org/bacpac/orderingframe.htm). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html Plate: 240 row: N column: 16 Seq primer: SP6 Class: BAC ends.

```

TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
PUBMED     10737800
COMMENT    Contact: Simpson A.J.G.
           Laboratory of Cancer Genetics
           Ludwig Institute for Cancer Research
           Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
           Brazil
           Tel: +55-11-2704922
           Fax: +55-11-2707001
           Email: asimpson@ludwig.org.br
           This sequence was derived from the FAPESP/LICR Human Cancer Genome
           Project. This entry can be seen in the following URL
           (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-RC3-HT0371-250)
           200-016-allst3-2000-02-25st4-1)
Seq primer: puc 18 forward
High quality sequence stop: 439.
Location/Qualifiers
FEATURES   1..439
            source
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /dev_stage="Adult"
            /clone_lib="HT0371"
            /note="Organ: head/neck; Vector: puc18; Site_1: SmaI;
            Site_2: SmaI; A mini-library was made by cloning products
            derived from ORESTES PCR (U.S. Letters Patent application
            No. 196,716 - Ludwig Institute for Cancer Research)
            profiles into the pUC 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
            low stringency conditions."
BASE COUNT 89 a 139 c 132 g 79 t
ORIGIN

Alignment Scores:
Pred. No.:      0.262      Length:      439
Score:          85.00      Matches:     23
Percent Similarity: 46.00%      Conservative: 0
Best Local Similarity: 46.00%      Mismatches: 3
Query Match:     53.46%      Indels:     24
DB:              10          Gaps:        2

JUNC_SEQ3_SEQ4 (1-30) x BE157169 (1-439)

Qy      5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
        |||||||
Db      167 CCTGCAGAGGTGGGAGGAGAGATGATGCCAGTATGCCAGGCCCTCACGGAAGGCTGC 108
        |||||||

Qy      16 -----GlnAsnGluAspLeu 20
        |||||||

Db      107 ATGCTGGGCTGGGAGGGGCCACCATCTGCTCTCTCTCCACAGATGAGGACTTG 48
        |||||||

Qy      21 GlyProAlaSerProLeuAspSerThrPhe 30
        |||||||

Db      47 GCCCCAGCCAGTCCCTTGGACAGCACCTTC 18
        |||||||

RESULT 4
LOCUS    CD558559
DEFINITION AGNCOURT_14477527 NIH_MGC_181 Homo sapiens cDNA clone
ACCESSION IMAGE:30396701 5', mRNA sequence.
VERSION    CD558559
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)

```

JOURNAL COMMENT

Unpublished
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabps-re@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: NDAM478 row: n column: 06
High quality sequence start: 19
High quality sequence stop: 537.

FEATURES

source

```

Location/Qualifiers
1..944
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="IMAGE:30396701"
/dev_stage="Unknown"
/lab_host="NIH_MGC_181"
/clone_lib="NIH_MGC_181"
/note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
(destroyed); Library is oligo-dr primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."
BASE COUNT 150 a 269 c 366 g 158 t 1 others
ORIGIN

```

Alignment Scores:

```

Pred. No.:      0.682      Length:      944
Score:          85.00      Matches:     23
Percent Similarity: 46.00%      Conservative: 0
Best Local Similarity: 46.00%      Mismatches: 3
Query Match:     53.46%      Indels:     24
DB:              14          Gaps:        2

```

JUNC_SEQ3_SEQ4 (1-30) x CD558559 (1-944)

```

Qy      5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
        |||||||
Db      193 CCTGCAGAGGTGGGAGGAGAGATGATGCCAGTATGCCAGGCCCTCACGGAAGGCTGC 252
        |||||||

Qy      16 -----GlnAsnGluAspLeu 20
        |||||||

Db      253 ATGCTGGGCTGGGAGGGGCCACCATCTGCTCTCTCTCCACAGATGAGGACTTG 312
        |||||||

Qy      21 GlyProAlaSerProLeuAspSerThrPhe 30
        |||||||

Db      313 GCCCCAGCCAGTCCCTTGGACAGCACCTTC 342
        |||||||

```

RESULT 5

LOCUS

BE157103 458 bp mRNA linear EST 21-JUN-2000
RC3-HT0371-180200-015-h07 HT0371 Homo sapiens cDNA, mRNA sequence.

DEFINITION

BE157103

ACCESSION

BE157103.1 GI:8619824

VERSION

EST.

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 458)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

FEATURES
source

Location/Qualifiers
1. .450
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT1160"

/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
84 a 127 c 150 g 89 t

BASE COUNT
ORIGIN

Alignment Scores:
Pred. No.: 1.17 Length: 450
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x BG988031 (1-450)

QY 1 AsPLySGlYcysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 259 GACAAGGGCTGCCCTCCGAGCAGAGAGCCCTCTGACGTC 303

RESULT 10

BM720186

LOCUS

BM720186 466 bp mRNA linear EST 01-MAR-2002
UI-E-E00-ahz-c-08-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone

DEFINITION UI-E-E00-ahz-c-08-0-UI 5', mRNA sequence.

ACCESSION BM720186

VERSION BM720186.1

KEYWORDS GI:19039265

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 466)

Bonaldo.M.F., Lennon.G. and Soares.M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bent-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Dr. M. Bento Soares, University of Iowa

Genetics (www.resgen.com).

Seq primer: M13 Reverse

Location/Qualifiers

1. .466
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-E00-ahz-c-08-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

FEATURES
source

/clone_lib="UI-E-E00"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-E00 is a cDNA library containing the following
tissue(s): fetal eye. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGGTATACC. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
BASE COUNT 105 a 114 c 151 g 93 t
ORIGIN

Alignment Scores:

Pred. No.: 1.23 Length: 466
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x BM720186 (1-466)

QY 1 AsPLySGlYcysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 18 GACAAGGGCTGCCCTCCGAGCAGAGAGCCCTCTGACGTC 62

RESULT 11

BI194790

LOCUS

BI194790 546 bp mRNA linear EST 10-JUL-2001
602948150F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5091396 5',
mRNA sequence.

DEFINITION

ACCESSION BI194790.1

VERSION GI:14649810

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 546)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@email.nih.gov

Tissue Procurement: AFCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCMI855 row: c column: 13

High quality sequence start: 10

High quality sequence stop: 495.

Location/Qualifiers

1. .546

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5091396"

/tissue_type="epithelioid carcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_42"

/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;

Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

FEATURES
source

following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).

BASE COUNT 110 a 147 c 183 g 106 t
ORIGIN

Alignment Scores:
Pred. No.: 1.49 Length: 546
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x BI194790 (1-546)

Qy 1 AsPLysGLyCysProAlaGLuInArgAlaSerProLeuThrSer 15
|||||
Db 223 GACAAGGCTGCCCGCGAGCAGAGCCGCTCTGAGCTCC 267

RESULT 12

AA443351

LOCUS

DEFINITION zw85b05.r1 Soares_total_fetus_Nb2HF8.9w Homo sapiens cDNA clone IMAGE:783729 5' similar to gb:M11730 ERBB-2 RECEPTOR PROTEIN-TYROSINE KINASE PRECURSOR (HUMAN); mRNA sequence.

ACCESSION AA443351

VERSION AA443351.1 GI:2156026

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.
Washu-Merck EST Project 1997

TITLE

JOURNAL

COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 360.

FEATURES

source

Location/Qualifiers
1..547
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5981347"
/db_xref="taxon:9606"
/clone="IMAGE:783729"
/dev_stage="8-9 weeks"
/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGCGCTTAATTTTTTTTTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 115 a 151 c 172 g 108 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 1.5 Length: 547
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 9 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x AA443351 (1-547)

Qy 1 AsPLysGLyCysProAlaGLuInArgAlaSerProLeuThrSer 15
|||||
Db 174 GACAAGGCTGCCCGCGAGCAGAGCCGCTCTGAGCTCC 218

RESULT 13

BM720098

LOCUS

DEFINITION UI-E-E00-ahz-c-05-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone UI-E-E00-ahz-c-05-0-UI 5', mRNA sequence.

ACCESSION BM720098

VERSION BM720098.1 GI:19039119

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Bonaldo, M.F., Lennon, G., and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL 97044477

MEDLINE 8889548

PUBMED

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

source

1..561
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-E00-ahz-c-05-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-E00"

/note="Organ: eye; Vector: pT7T3-pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-E00 is a cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCGTATACC. This library

was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 97 a 171 c 179 g 113 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1.54 Length: 561
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x BM720098 (1-561)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 400 GACAAGGGCTGCCCGCAGCAGAGCCGCTCTGACGTC 444

RESULT 14

BM083444

LOCUS

DEFINITION BM083444 607 bp mRNA linear EST 16-NOV-2001
IMAGE:6-2000/sjp459bdf42.x2 NIH_MGC_39 Homo sapiens cDNA clone
IMAGE:3610674 5', mRNA sequence.

ACCESSION

BM083444

VERSION

BM083444.1 GI:16951075

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Kale, P.I.; Harsch, T.J.; Folta, P.A.; Nelson, D.O.; Sanders, C.G. and

Prange, C.K.

The I.M.A.G.E. Consortium quality control effort: clone

resequencing for verification

Other_ESTs: BB615590

Contact: Prange CK

The I.M.A.G.E. Consortium

Lawrence Livermore National Laboratory

Livermore, CA, USA

Email: help@image.llnl.gov

This read has been verified (found to hit its original self in the

correct orientation), as part of the I.M.A.G.E. Consortium quality

control effort. High quality sequence is defined as having 100 or

more base pairs with a phred quality value of 20 or greater, where

a sliding window of 4 base pairs with a phred quality value of 15

or greater marks the beginning and end of the sequence. For

information on obtaining this clone, please contact

info@image.llnl.gov.

Plate: LLCM267 row: b column: 19

Seq primer: -2lm13

High quality sequence stop: 607.

Location/Qualifiers

1..607

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3610674"

/tissue_type="adenocarcinoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_39"

/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;

Site:2: EcoRI; cDNA made by oligo-dt priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed

by Ling Hong in the laboratory of Gerald M. Rubin

(University of California, Berkeley) using ZAP-cDNA

synthesis kit (Stratagene) and Superscript II RT (Life

Technologies)."

BASE COUNT

ORIGIN

118 a 171 c 195 g 121 t 2 others

Alignment Scores:

Pred. No.: 1.7 Length: 607
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x BM083444 (1-607)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 292 GACAAGGGCTGCCCGCAGCAGAGCCGCTCTGACGTC 336

RESULT 15

AW410534

LOCUS

DEFINITION AW410534 613 bp mRNA linear EST 29-JUN-2000
fh06h06.x1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2961635 5',
mRNA sequence.

ACCESSION

AW410534

VERSION

AW410534.1 GI:6936075

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgc.nci.nih.gov/.

1 (bases 1 to 613)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/dbrrp/image/image.html

Plate: LLCM56 row: 0 column: 12

Seq primer: -2lm13 forward primer (ABI).

Location/Qualifiers

1..613

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2961635"

/tissue_type="rhabdomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_17"

/note="Organ: muscle; Vector: pOTB7; Site:1: EcoRI;

Site:2: XhoI; cDNA made by oligo-dt priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 109 a 205 c 185 g 113 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1.73 Length: 613

Score: 80.00 Matches: 15

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 50.31% Indels: 0

DB: 9 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x AW410534 (1-613)

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db 557 GACAAAGGGCTGCCCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCC 601

Search completed: October 16, 2003, 17:04:02
Job time : 237.794 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 11:08:07 : Search time 31.0111 Seconds
(without alignments)
2540.503 Million cell updates/sec

Title: JUNC_SEQ3_SEQ4

Perfect score: 159

Sequence: 1 DKGPCAQRASPLTSQNEGLGASPLDSTF 30

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool_p/HOLLERAN80/funat_15102003_131915_20662/app_query.fasta_1.4685
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=HOLLERAN80 -CGN_1_1_397 -runat_15102003_131915_20662
-NCPU=6 -ICPU=3 -NO_WMAP -LANGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCIT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCITUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	50.3	201	13	US-10-109-213-3
Sequence 3, Appli					

2	80	50.3	3765	12	US-10-207-498-5	Sequence 5, Appli
3	80	50.3	3768	9	US-09-811-123-8	Sequence 8, Appli
4	80	50.3	3768	9	US-09-811-115-2	Sequence 2, Appli
5	80	50.3	3768	10	US-09-854-356-9	Sequence 9, Appli
6	80	50.3	3768	10	US-09-930-125-1	Sequence 1, Appli
7	80	50.3	3768	12	US-10-313-644-1	Sequence 1, Appli
8	80	50.3	4473	11	US-09-441-411-5	Sequence 5, Appli
9	80	50.3	4473	12	US-10-101-510-81	Sequence 81, Appli
10	80	50.3	4473	14	US-10-146-473-32	Sequence 32, Appli
11	80	50.3	4473	14	US-10-207-655-44	Sequence 44, Appli
12	80	50.3	4530	10	US-09-877-177-11	Sequence 11, Appli
13	80	50.3	4530	12	US-10-087-926A-119	Sequence 119, App
14	80	50.3	4530	12	US-10-101-510-124	Sequence 124, App
15	80	50.3	4530	12	US-10-338-730-1	Sequence 1, Appli
16	80	50.3	4530	14	US-10-177-293-125	Sequence 125, App
17	80	50.3	4543	10	US-09-769-508-1	Sequence 1, Appli
18	80	50.3	4606	12	US-09-971-392-70	Sequence 70, Appl
19	80	50.3	4642	14	US-10-198-846-10896	Sequence 10896, A
20	80	50.3	9274	9	US-09-811-123-7	Sequence 7, Appli
21	80	50.3	9274	9	US-09-811-115-1	Sequence 1, Appli
22	79	49.7	1115	14	US-10-102-806-165	Sequence 165, App
23	79	49.7	1713	12	US-10-378-393-14	Sequence 14, Appl
24	79	49.7	1755	10	US-09-930-125-6	Sequence 6, Appli
25	79	49.7	1767	10	US-09-930-125-4	Sequence 4, Appli
26	79	49.7	1773	10	US-09-930-125-7	Sequence 7, Appli
27	79	49.7	1806	10	US-09-930-125-5	Sequence 5, Appli
28	79	49.7	2411	12	US-10-378-393-10	Sequence 10, Appl
29	74	46.5	3771	10	US-09-854-356-11	Sequence 11, Appl
30	74	46.5	3955	10	US-09-870-759-117	Sequence 117, App
31	74	46.5	3955	10	US-09-854-356-10	Sequence 10, Appl
32	74	46.5	3955	12	US-09-751-708A-117	Sequence 117, App
33	62	39.0	14427	14	US-10-156-761-1540	Sequence 1540, Ap
34	62	39.0	23432	9	US-09-764-869-1332	Sequence 1332, Ap
35	62	39.0	23432	14	US-10-091-504-1332	Sequence 1332, Ap
36	62	39.0	9025608	14	US-10-156-761-1	Sequence 1, Appli
37	59	37.1	111822	13	US-10-094-989-3	Sequence 3, Appli
38	58.5	36.8	633	11	US-09-788-188-12	Sequence 12, Appl
39	58.5	36.8	634	8	US-08-450-842-1	Sequence 1, Appli
40	58.5	36.8	1404	14	US-10-155-785-13	Sequence 13, Appl
41	57.5	36.2	363	11	US-09-918-995-28706	Sequence 28706, A
42	57	35.8	297	10	US-09-783-590-11498	Sequence 11498, A
43	57	35.8	1938	12	US-10-422-264-25	Sequence 25, Appl
44	57	35.8	2316	12	US-10-422-264-27	Sequence 27, Appl
45	57	35.8	2604	12	US-10-422-264-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1

US-10-109-213-3
: Sequence 3, Application US/10109213
: Publication No. US20020186670A1
: GENERAL INFORMATION:
: APPLICANT: Ecker, David J.
: TITLE OF INVENTION: Identification Of Disease Predictive Nucleic Acids
: FILE REFERENCE: IBIS0009
: CURRENT APPLICATION NUMBER: US/10/109,213
: CURRENT FILING DATE: 2002-03-27
: PRIOR APPLICATION NUMBER: US/09/200,355
: PRIOR FILING DATE: 1998-11-25
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 3
: LENGTH: 201
: TYPE: RNA
: ORGANISM: Homo sapiens
US-10-109-213-3

Alignment Scores:	0.000724	201
Pred. No.:	80.00	15
Score:	100.00%	0
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	0

Query Match: 50.31% Indels: 0
DB: 13 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-10-109-213-3 (1-201)

QY 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
DB 116 GACAAGGGCGUGCCCGCCGAGCAGAGAGCCGCCUCGACGUC 160

RESULT 2

US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; FILE REFERENCE: 30448.103-US-01
; CURRENT APPLICATION NUMBER: US/10/207,498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5

Alignment Scores:
Pred. No.: 0.0177 Length: 3765
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-10-207-498-5 (1-3765)

QY 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
DB 1915 GACAAGGGTGTCCCGCCGAGCAGAGAGCCGCCCTTGACGTCC 1959

RESULT 3

US-09-811-123-8
; Sequence 8, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-811-123-8

Alignment Scores:

Pred. No.: 0.0177 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 9 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-09-811-123-8 (1-3768)

QY 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
DB 1915 GACAAGGGTGTCCCGCCGAGCAGAGAGCCGCCCTTGACGTCC 1959

RESULT 4

US-09-811-115-2
; Sequence 2, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred. No.: 0.0177 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 9 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-09-811-115-2 (1-3768)

QY 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
DB 1915 GACAAGGGTGTCCCGCCGAGCAGAGAGCCGCCCTTGACGTCC 1959

RESULT 5

US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768

```

Alignment Scores:
  Pred. No.:      0.0214      Length:      4473
  Score:          80.00      Matches:      15
  Percent Similarity: 100.00%      Conservative: 0
  Best Local Similarity: 100.00%      Mismatches: 0
  Query Match:      50.31%      Indels:      0
  DB:              11      Gaps:      0

JUNC_SEQ3_SEQ4 (1-30) x US-09-441-411-5 (1-4473)

```

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGGCTGCCCCGGCGAGCAGAGAGCCAGCCCTCTGAGCTCC 2133

RESULT 9

US-10-101-510-81

; Sequence 81, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-81

Alignment Scores:
Pred. No.: 0.0214 Length: 4473
Score: 80.00 Matches: 15
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 50.31%
Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-10-101-510-81 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGGCTGCCCCGGCGAGCAGAGAGCCAGCCCTCTGAGCTCC 2133

RESULT 10

US-10-146-473-32

; Sequence 32, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-473-32

Alignment Scores:
Pred. No.: 0.0214 Length: 4473
Score: 80.00 Matches: 15
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 50.31%
Indels: 0
DB: 14 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-10-146-473-32 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGGCTGCCCCGGCGAGCAGAGAGCCAGCCCTCTGAGCTCC 2133

RESULT 11

US-10-207-655-44
; Sequence 44, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069, 401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-207-655-44

Alignment Scores:
Pred. No.: 0.0214 Length: 4473
Score: 80.00 Matches: 15
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 50.31%
Indels: 0
DB: 14 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-10-207-655-44 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGGCTGCCCCGGCGAGCAGAGAGCCAGCCCTCTGAGCTCC 2133

RESULT 12

US-09-877-177-11

; Sequence 11, Application US/09877177
; Publication No. US20020192652A1
; GENERAL INFORMATION:
; APPLICANT: Peter V. Danenberg et al.
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-New Gene Expression
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177-11

Alignment Scores:
Pred. No.: 0.0217 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 50.31%
Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-09-877-177-11 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2065 GACAAGGGCTGCCCCGGCGAGCAGAGAGCCAGCCCTCTGAGCTCC 2109

RESULT 13

US-10-007-926A-119

```

: Sequence 119, Application US/10007926A
: Publication No. US20030143539A1
: GENERAL INFORMATION:
: APPLICANT: BERTUCCI, FRANCOIS
: APPLICANT: HOULGATTE, REMI
: APPLICANT: BIRNBAUM, DANIEL
: APPLICANT: NGUYEN, CATHERINE
: APPLICANT: VIENS, PATRICE
: APPLICANT: FERT, VINCENT
: TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
: TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
: FILE REFERENCE: 1546-R-00
: CURRENT APPLICATION NUMBER: US/10/007,926A
: CURRENT FILING DATE: 2001-12-07
: PRIOR APPLICATION NUMBER: 60/254,090
: PRIOR FILING DATE: 2000-12-08
: NUMBER OF SEQ ID NOS: 468
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 119
: LENGTH: 4530
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: v-erb-b2 avian erythroblastic leukemia viral
: OTHER INFORMATION: oncogene homolog 2 (neuro/glioblastoma derived
: OTHER INFORMATION: oncogene homolog) (ERBB2) gene.
US-10-007-926A-119

Alignment Scores:
Pred. No.: 0.0217 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-10-007-926A-119 (1-4530)
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db 2065 GACAAGGGTGTCCCCCGCAGAGAGCCAGCCCTCTGACGTC 2109

RESULT 14
US-10-101-510-124
: Sequence 124, Application US/10101510
: Publication No. US20030148295A1
: GENERAL INFORMATION:
: APPLICANT: WAN, JACKSON
: APPLICANT: WANG, YIXIN
: TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
: FILE REFERENCE: 15117.0012
: CURRENT APPLICATION NUMBER: US/10/101,510
: CURRENT FILING DATE: 2002-03-20
: PRIOR APPLICATION NUMBER: 60/276,947
: PRIOR FILING DATE: 2001-03-20
: NUMBER OF SEQ ID NOS: 805
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 124
: LENGTH: 4530
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-101-510-124

Alignment Scores:
Pred. No.: 0.0217 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-10-101-510-124 (1-4530)

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: October 15, 2003, 22:55:19 ; Search time 7.60488 Seconds
(without alignments)
1741.185 Million cell updates/sec

Title: JUNC_SEQ3_SEQ4
Perfect score: 159
Sequence: 1 DKGPCAEGRASPLTSQNEGLGPASPLDSTF 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/HOLLERAN480/runat_15102003_131913_20570/app_query.fasta_1.4685
-DB=Issued_Patents_NA -QMT=Fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTEMP=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGN_1_1_403=runat_15102003_131913_20570 -NCPU=6 -ICPU=3
-DEV_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	52.8	2385	2	US-09-146-283-3
2	84	52.8	2385	3	US-08-579-823A-3
3	84	52.8	2385	3	US-09-344-195-3
4	80	50.3	153	3	US-08-776-251-3
5	80	50.3	201	4	US-09-200-355-3
6	80	50.3	816	3	US-08-776-251-10
7	80	50.3	3768	2	US-08-625-101-1
8	80	50.3	3768	2	US-08-356-786-1
9	80	50.3	4473	2	US-09-048-804-1
10	80	50.3	4473	3	US-09-056-105-26
11	80	50.3	4530	1	US-08-229-515A-9
12	80	50.3	4530	1	US-08-645-865-9

13	80	50.3	4530	4	US-09-167-322-4	Sequence 4, Appli
14	80	50.3	4530	4	US-09-527-487-1	Sequence 1, Appli
15	80	50.3	4530	4	US-09-877-177A-11	Sequence 11, Appl
16	74	46.5	3955	1	US-08-229-515A-14	Sequence 14, Appl
17	74	46.5	3955	1	US-08-645-865-14	Sequence 14, Appl
c 18	59	37.1	111282	4	US-09-754-250-3	Sequence 3, Appli
c 19	58.5	36.8	634	1	US-08-451-947-1	Sequence 1, Appli
c 20	58.5	36.8	634	2	US-08-424-826A-1	Sequence 1, Appli
c 21	58.5	36.8	634	3	US-08-928-694-1	Sequence 1, Appli
c 22	58.5	36.8	634	4	US-08-450-842-1	Sequence 1, Appli
c 23	58.5	36.8	634	4	US-08-451-390-1	Sequence 1, Appli
c 24	58.5	36.8	634	5	PCT-US91-06950-1	Sequence 1, Appli
c 25	58.5	36.8	1404	1	US-07-796-105-22	Sequence 22, Appl
c 26	58.5	36.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c 27	58.5	36.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 28	57	35.8	1938	4	US-09-547-435-25	Sequence 25, Appl
c 29	57	35.8	2316	4	US-09-547-435-27	Sequence 27, Appl
c 30	57	35.8	2604	4	US-09-547-435-23	Sequence 23, Appl
c 31	57	35.8	3384	4	US-09-547-435-29	Sequence 29, Appl
c 32	55.5	34.9	2691	4	US-09-020-743-1	Sequence 1, Appli
c 33	55	34.6	960	3	US-08-651-136C-1	Sequence 1, Appli
c 34	55	34.6	960	4	US-09-229-911A-1	Sequence 1, Appli
c 35	55	34.6	4092	3	US-09-306-595C-5	Sequence 5, Appli
c 36	55	34.6	4092	4	US-09-925-388-5	Sequence 5, Appli
c 37	55	34.6	5046	4	US-09-548-938A-5	Sequence 5, Appli
c 38	54.5	34.3	824	4	US-09-312-283C-369	Sequence 369, App
c 39	53.5	33.6	685	1	US-08-451-947-7	Sequence 7, Appli
c 40	53.5	33.6	685	2	US-08-424-826A-7	Sequence 7, Appli
c 41	53.5	33.6	685	3	US-08-928-694-7	Sequence 7, Appli
c 42	53.5	33.6	685	4	US-08-450-842-7	Sequence 7, Appli
c 43	53.5	33.6	685	4	US-08-451-390-7	Sequence 7, Appli
c 44	53.5	33.6	685	5	PCT-US91-06950-7	Sequence 7, Appli
c 45	53.5	33.6	1190	1	US-08-451-947-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Relner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs

TYPE:	nucleic acid
STRANDEDNESS:	double
TOPOLOGY:	linear
MOLECULE TYPE:	DNA (genomic)
HYPOTHETICAL:	NO
ANTI-SENSE:	NO
ORIGINAL SOURCE:	
ORGANISM:	homo sapiens
INDIVIDUAL ISOLATE:	GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3	
Alignment Scores:	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-09-146-283-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	0.0053
Score:	84.00
Percent Similarity:	70.00%
Best Local Similarity:	70.00%
Query Match:	52.83%
DB:	2
JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)	
Pred. No.:	

Pred. No.: 0.0053 Length: 2385
Score: 84.00 Matches: 21
Percent Similarity: 70.00% Conservative: 0
Best Local Similarity: 70.00% Mismatches: 5
Query Match: 52.83% Indels: 4
DB: 3 Gaps: 2

JUNC_SEQ3_SEQ4 (1-30) x US-09-344-195-3 (1-2385)

Qy 1 AsplysGlyCysProAlaGluClnArgAlaSerProLeuThrSerGlnAsnGluAspleu 20
|||||
Db 1925 GACAAGGCTGCCCGCGAGAGAGCCAGCCCTCTGACGTCCCTCGAG----- 1975

Qy 21 GlyProAla---SerProLeuAspSerThr 29
|||||

Db 1976 GCACCCCGCGCTCGCCCGAGCCCGCCAGCAC 2005

RESULT 4

US-08-776-251-3
; Sequence 3, Application US/08776251
; Patent No. 6025340
; GENERAL INFORMATION:
; APPLICANT: Springer, Caroline J
; APPLICANT: Marais, Richard
; TITLE OF INVENTION: Surface expression of enzyme in gene directed prodrgug therapy
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6025340th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,251
; FILING DATE: 31-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/01782
; FILING DATE: 27-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9415167.7
; FILING DATE: 27-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 620-20
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 153 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-776-251-3

Alignment Scores:
Pred. No.: 0.000654 Length: 153
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-08-776-251-3 (1-153)

Qy 1 AsplysGlyCysProAlaGluClnArgAlaSerProLeuThrSer 15
|||||
Db 10 GACAAGGCTGCCCGCGAGAGAGCCAGCCCTCTGACGTCC 54

RESULT 5
US-09-200-355-3
; Sequence 3, Application US/09200355
; Patent No. 6451524
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; TITLE OF INVENTION: Identification Of Disease Predictive Nucleic Acids
; FILE REFERENCE: IBIS0009
; CURRENT APPLICATION NUMBER: US/09/200,355
; CURRENT FILING DATE: 1998-11-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 201
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-200-355-3

Alignment Scores:
Pred. No.: 0.000922 Length: 201
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-09-200-355-3 (1-201)

Qy 1 AsplysGlyCysProAlaGluClnArgAlaSerProLeuThrSer 15
|||||
Db 116 GACAAGGCGCGCCCGAGAGAGCCAGCCCGCUGAGGUCC 160

RESULT 6

US-08-776-251-10
; Sequence 10, Application US/08776251
; Patent No. 6025340
; GENERAL INFORMATION:
; APPLICANT: Springer, Caroline J
; APPLICANT: Marais, Richard
; TITLE OF INVENTION: Surface expression of enzyme in gene directed prodrgug there
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6025340th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,251
; FILING DATE: 31-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/01782
; FILING DATE: 27-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9415167.7
; FILING DATE: 27-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 620-20
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-776-251-10

Alignment Scores:
 Pred. No.: 0.00536 Length: 816
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 3 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-08-776-251-10 (1-816)

QY 1 AsPLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
 ||||||||||||||||||||||||||||||||||||||||

Db 93 GACAAGGCTGCCCGCCGAGAGAGAGCCGCTCTGACGTCC 137

RESULT 7
 US-08-625-101-1
 ; Sequence 1, Application US/08625101
 ; Patent No. 5869445
 ; GENERAL INFORMATION:
 ; APPLICANT: Cheever, Martin A.
 ; APPLICANT: Disis, Mary L.
 ; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
 ; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
 ; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
 ; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED and BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98104-7092
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/625,101
 ; FILING DATE: 01-APR-1996
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sharkey, Richard G.
 ; REGISTRATION NUMBER: 32,629
 ; REFERENCE/DOCKET NUMBER: 920010.448C7
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3768 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..3765
 ; US-08-625-101-1

Alignment Scores:
 Pred. No.: 0.0366 Length: 3768
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 2 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-08-625-101-1 (1-3768)

QY 1 AsPLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
 ||||||||||||||||||||||||||||||||||||||||

Db 1915 GACAAGGGCTGCCCGCCGAGAGAGAGCCGCTCTGACGTCC 1959

RESULT 8
 US-08-356-786-1
 ; Sequence 1, Application US/08356786
 ; Patent No. 5877305
 ; GENERAL INFORMATION:
 ; APPLICANT: Huston, James S.
 ; APPLICANT: Oppermann, Hermann
 ; APPLICANT: Houston, L. L.
 ; APPLICANT: Ring, David B.
 ; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
 ; TITLE OF INVENTION: Marker
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
 ; STREET: Exchange Place, 53 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/356,786
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/831,967
 ; FILING DATE: 06-FEB-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pitcher, Edmund R.
 ; REGISTRATION NUMBER: 27,829
 ; REFERENCE/DOCKET NUMBER: CRP-053
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 248-7000
 ; TELEFAX: (617) 248-7100
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3768 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..3768
 ; OTHER INFORMATION: /note= "product = "cerB-b2"
 ; US-08-356-786-1

Alignment Scores:
 Pred. No.: 0.0366 Length: 3768
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 2 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-08-356-786-1 (1-3768)

QY 1 AsPLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
 ||||||||||||||||||||||||||||||||||||||||

Db 1915 GACAAGGGCTGCCCGCCGAGAGAGAGCCGCTCTGACGTCC 1959

RESULT 9
 US-09-048-804-1
 ; Sequence 1, Application US/09048804
 ; Patent No. 5968748
 ; GENERAL INFORMATION:
 ; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A-9

Alignment Scores:
Pred. No.:      0.0462          Length:      4530
Score:          80.00           Matches:     15
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:   0
Query Match:      50.31%       Indels:      0
DB:               1           Gaps:        0

JUNC_SEQ3_SEQ4 (1-30) x US-08-229-515A-9 (1-4530)

Qy      1 AsplysGlycysProAlaGluglnArgAlaSerProLeuthrSer 15
         |||
Db      2065 GACAAGGGCTGCCCGCCGACGACAGAGCCAGCCCTCTGCAGTCC 2109

RESULT 12
US-08-645-865-9
; Sequence 9, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZI TUDDIN
```

```

; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-645-865-9

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-08-645-865-9 (1-4530)
QY 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db 2065 GACAGGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTCC 2109

RESULT 13
US-09-167-322-4
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322

```

```

; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
; US-09-167-322-4

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-09-167-322-4 (1-4530)
QY 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db 2065 GACAGGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTCC 2109

RESULT 14
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
; US-09-527-487-1

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-09-527-487-1 (1-4530)
QY 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db 2065 GACAGGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTCC 2109

RESULT 15
US-09-877-177A-11
; Sequence 11, Application US/09877177A

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

and is derived by analysis of the total score distribution.

SUMMARIES

OM protein - nucleic search, using frame_plus_p2n model									
Run on:	October 15, 2003, 21:06:28	Search time	28.4941 Seconds						
(without alignments)									
2842.104 Million cell updates/sec									
Title:	JUNC_SEQ3_SEQ4								
Perfect score:	159								
Sequence:	1 DKGCPAEQRAASPLTSQNEGLGAPLPLDSTF 30								
Scoring table:	BLOSUM62								
	Xgapop 10.0 , Xgapext 0.5								
	Ygapop 10.0 , Ygapext 0.5								
	Fgapop 6.0 , Fgapext 7.0								
	Delop 6.0 , Delext 7.0								
Searched:	2552756 seqs, 1349719017 residues								
Total number of hits satisfying chosen parameters:	5105512								
Minimum DB seq length:	0								
Maximum DB seq length:	2000000000								
Post-processing:	Minimum Match 0%								
	Maximum Match 100%								
	Listing first 45 summaries								
Command line parameters:									
-MODEL=frame+p2n.model -DEV=xlp									
-Q/cygn2_1/USPTO_spool_p/HOLLERAN480/runat_15102003_131912_20526/app_query.fasta_1.4685									
-DB=N_Geneseq_19Jun03 -Qfmt=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0									
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi									
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15									
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000									
-USER=HOLLERAN480_@CN_1_1_1758_@runat_15102003_131912_20526 -NCPU=6 -ICPU=3									
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG									
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6									
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7									
Database :	N_Geneseq_19Jun03.*								
1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*								
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*								
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*								
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*								
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*								
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*								
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*								
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*								
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*								
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*								
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*								
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*								
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*								
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*								
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*								
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*								
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*								
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*								
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*								
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*								
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*								
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*								
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*								
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*								
25:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*								
26:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2004.DAT.*								
27:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2005.DAT.*								
28:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2006.DAT.*								
29:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2007.DAT.*								
30:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2008.DAT.*								
31:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2009.DAT.*								
32:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2010.DAT.*								
33:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2011.DAT.*								
34:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2012.DAT.*								
35:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2013.DAT.*								
36:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2014.DAT.*								
37:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2015.DAT.*								
38:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2016.DAT.*								
39:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2017.DAT.*								
40:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2018.DAT.*								
41:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2019.DAT.*								
42:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2020.DAT.*								
43:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2021.DAT.*								
44:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2022.DAT.*								
45:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2023.DAT.*								

ALIGNMENTS

RESULT 1	
ABA92252	ID ABA92252 standard; cdna; 2763 BP.
XX	AC ABA92252;
XX	DT 17-JUN-2002 (first entry)
XX	DE Mouse Her-2/neu extracellular-phosphorylation domain fusion cdna.
XX	KW Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
KW	receptor; mouse; gene therapy; gene; ss.
XX	OS Mus musculus.
XX	Key Location/Qualifiers
FT	CDS 1..2763

QY 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
 Db 1918 GAACGAGGCTGCCAGCAGAGCAGAGCCAGCCAGTCTCAGAACGAGGACTTAA 1977
 QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
 Db 1978 GGCCCTCCAGCCCATGGAGACAGCACCCTC 2007

RESULT 3

AAAT72725
 ID AAAT72725 standard; cDNA; 2385 BP.
 AC AAAT72725;
 DT 17-SEP-1997 (first entry)
 DE Her2-GM-CSF immunostimulant fusion protein DNA.
 KW Her2-GM-CSF; granulocyte macrophage colony stimulating factor;
 KW growth factor receptor; oncogene; immunostimulant; cancer;
 KW therapy; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 11..2359
 FT CDS /*tag= a
 FT /product= GM-CSF-Her2 fusion protein
 FT mRNA 11..1969
 FT /*tag= b
 FT /product= Her2
 FT mRNA 1970..1975
 FT /*tag= c
 FT /product= Leu-Glu linker
 FT mRNA 1976..2359
 FT /*tag= d
 FT /product= GM-CSF
 XX
 XX WO9724438-A1.
 XX
 XX 10-JUL-1997.
 XX
 XX 23-DEC-1996; 96WO-US20241.
 XX
 XX 28-DEC-1995; 95US-0579823.
 XX
 XX (ACTI-) ACTIVATED CELL THERAPY INC.
 XX
 XX Laus R, Ruegg CL, Wu H;
 XX WPI; 1997-363674/33.
 DR P-PSDB; AAW19764.
 XX
 XX Potent APC that activates T-cells to give multivalent cellular
 PT immune response - can also induce a cytotoxic T-cell response in a
 PT vertebrate subject
 PT
 XX Disclosure; Fig 8; 45pp; English.
 XX
 XX A nucleic acid molecule (AAAT72725) codes for a fusion protein
 CC (AAW19764) comprising granulocyte-macrophage colony stimulating
 CC factor (GM-CSF) and Her2, a growth factor receptor that is
 CC over-expressed in breast and ovarian cancer cells. It was
 CC prepd. by PCR amplification of Her2 cDNA from a breast cancer
 CC cell line and fusion to GM-CSF cDNA. Fusion expression vectors can
 CC be used to transfect mammalian and insect cells. The Her2-GM-CSF
 CC fusion protein is used to generate anti-Her2 immunity. Tumour
 CC cells are eliminated by cytotoxic T lymphocytes activated in vivo
 CC or in vitro by exposure to antigen-presenting cells exposed to the
 CC fusion protein.
 XX
 XX Sequence 2385 BP; 488 A; 780 C; 677 G; 440 T; 0 other;

Alignment Scores:
 Pred. No.: 0.0382 Length: 2385
 Score: 84.00 Matches: 21
 Percent Similarity: 70.00% Conservative: 0
 Best Local Similarity: 70.00% Mismatches: 5
 Query Match: 52.83% Indels: 4
 DB: 18 Gaps: 2

JUNC_SEQ3_SEQ4 (1-30) x AAAT72725 (1-2385)

QY 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
 Db 1925 GACAAGGGCTGCCCGCCGAGAGAGCCAGCCCTCTGACGTCCCTCGAG----- 1975

QY 21 GlyProAla---SerProLeuAspSerThr 29
 Db 1976 GCACCGCGCGCTCGCCCGAGCCAGCCAGCACA 2005

RESULT 4
 AAAT4815
 ID AAAT4815 standard; cDNA; 69 BP.
 AC AAAT4815;
 XX
 XX 08-AUG-2000 (first entry)
 DE DNA encoding a wild type erbB-2 receptor protein fragment.
 XX
 KW SPLICE erbB-2 receptor protein; cell transformation disorder; cancer;
 KW tumor cell proliferation; tissue degeneration; arthropathy;
 KW bone resorption; inflammatory disease; degenerative disorder;
 KW wound healing; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200020579-A1.
 XX
 XX 13-APR-2000.
 XX
 XX 01-OCT-1999; 99WO-CA00912.
 XX
 XX 02-OCT-1998; 98US-0165192.
 XX
 XX (UYMC-) UNIV MCMASTER.
 XX
 XX Muller WJ, Siegel PM;
 XX WPI; 2000-303768/26.
 DR P-PSDB; AAY84681.
 XX
 XX Nucleic acid encoding an erbB 2 receptor protein designated SPLICE
 PT erbB-2, inhibitors of the protein are useful for treatment of cancer -
 PT
 XX Example 2; Fig 3B; 60pp; English.
 XX
 XX The present sequence encodes a peptide of wild type erbB-2 receptor
 CC protein, where the splicing event occurs. An alternatively spliced
 CC variant of erbB-2 receptor protein, designated SPLICE erbB-2 receptor
 CC protein, has an in-frame deletion of 16 amino acids, 2 of which are
 CC conserved cysteine residues, compared to the unspliced protein. The
 CC erbB-2 polynucleotide is used to construct probes for detecting
 CC disorders of cell transformation such as cancer. Antibodies to the
 CC protein may be used to detect SPLICE erbB-2 in a sample. Agents
 CC (e.g. antisense oligonucleotides) which inhibit the expression of
 CC SPLICE erbB-2 are useful for reducing tumor cell proliferation and
 CC treating cancer. Substances which stimulate SPLICE erbB-2 are useful
 CC for treating conditions involving damaged cells including conditions in
 CC which degeneration of tissue occurs, such as arthropathy, bone
 CC resorption, inflammatory diseases, degenerative disorders of the
 CC central nervous system and wound healing.
 XX
 XX Sequence 69 BP; 13 A; 26 C; 20 G; 10 T; 0 other;

```

Alignment Scores:
Pred. No.:      0.00195      Length:      69
Score:          80.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      50.31%      Indels:      0
DB:              21      Gaps:      0

JUNC_SEQ3_SEQ4 (1-30) x AAA14815 (1-69)

QY      1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db      25 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTC 69

RESULT 5
AAA62125
ID      AAA62125 standard; RNA; 201 BP.
XX
AC      AAA62125;
XX
DT      20-JUN-2001 (first entry)
XX
DE      Her2 normal form transcript.
XX
KW      Her2; disease prediction; breast cancer; human; tyrosine kinase; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      stem_loop      Location/Qualifiers
FT      stem_loop      8..18      /*tag= a
FT      stem_loop      37..56      /*tag= b
FT      misc_binding    58..59      /*tag= c
FT      stem_loop      /bound_moiety= "Her2 normal form transcript"
FT      /note= "Forms double-stranded region with bases 74-75
FT      of Her2 normal form transcript"
FT      stem_loop      62..71      /*tag= d
FT      misc_binding    74..75      /*tag= e
FT      /bound_moiety= "Her2 normal form transcript"
FT      /note= "Forms double-stranded region with bases 58-59
FT      of Her2 normal form transcript"
FT      misc_binding    78..81      /*tag= f
FT      /bound_moiety= "Her2 normal form transcript"
FT      /note= "Forms double-stranded region with bases 121-124
FT      of Her2 normal form transcript"
FT      misc_binding    84..90      /*tag= g
FT      /bound_moiety= "Her2 normal form transcript bases"
FT      /note= "Forms double-stranded region with bases 111-117
FT      of Her2 normal form transcript"
FT      stem_loop      94..107      /*tag= h
FT      misc_binding    111..117      /*tag= i
FT      /bound_moiety= "her2 normal form transcript"
FT      /note= "Forms double-stranded region with bases 84-90
FT      of Her2 normal form transcript"
FT      misc_binding    121..124      /*tag= j
FT      /bound_moiety= "Her2 normal form transcript"
FT      /note= "Forms double-stranded region with bases 78-81
FT      of Her2 normal form transcript"
FT      stem_loop      126..132      /*tag= k
FT      misc_binding    137..152      /*tag= l
FT      /bound_moiety= "Her2 normal form transcript bases"

/note= "Forms double-stranded region with bases 172-187
of Her2 normal form transcript"
/*tag= m
/*tag= n
/note= "her2 normal form transcript bases"
/bound_moiety= "her2 normal form transcript bases"
/note= "Forms double-stranded region with bases 137-152
of Her2 normal form transcript"

stem_loop
misc_binding
/*tag= m
/*tag= n
/note= "Forms double-stranded region with bases 172-187
of Her2 normal form transcript"

WO2000311110-A1.
XX
XX      02-JUN-2000.
PD
XX
XX      22-NOV-1999; 99WO-US27710.
PF
XX
XX      25-NOV-1998; 98US-0110024.
PR
XX      25-NOV-1998; 98US-0200355.
PR
XX
XX      (ISIS-) ISIS PHARM INC.
PA
XX
XX      Ecker DJ;
PI
XX
XX      WPI; 2000-400027/34.
DR
XX
XX      Identifying a target nucleic acid sequence predictive of preselected
PT      disease states such as a cancerous state, by comparing members of a set
PT      of mRNA molecules, from a common gene, containing different sequences
PT      and structures -
XX
XX      Example 2; Fig 3; 38pp; English.
PS
XX
XX      The HER2 proto-oncogene encodes a tyrosine kinase receptor, which
CC      functions in cellular signal transduction. The HER2 protein is
CC      implicated in breast cancer. The HER2 receptor mRNA exists in at least
CC      two forms: a normal form and a truncated form. The present sequence is
CC      the normal form RNA of HER2. The truncated form RNA sequence of HER2
CC      (AAA62126) encodes a protein associated with increased resistance to the
CC      growth inhibiting effects of a monoclonal antibody, Herceptin, used in
CC      cancer treatment. The truncated form of the transcript contains unique
CC      structures not found in the normal form. The present sequence is
CC      predictive of Herceptin-resistant cancer and detection of the present
CC      sequence may therefore be used as a method of diagnosing breast cancer.
CC      Other diseases which may be identified by using a similar method to
CC      detect other RNA molecules are hyperproliferative conditions, Lupus
CC      erythematosus, psoriasis, inflammation, cardiovascular disease, pain,
CC      arthritis, obesity, trauma, Huntington's disease or neurological
CC      disorders.
XX
XX      Sequence 201 BP; 34 A; 67 C; 56 G; 44 U; 0 other;
SQ
XX
XX      Alignment Scores:
XX      Pred. No.:      0.00714      Length:      201
XX      Score:          80.00      Matches:      15
XX      Percent Similarity: 100.00%      Conservative: 0
XX      Best Local Similarity: 100.00%      Mismatches: 0
XX      Query Match:      50.31%      Indels:      0
XX      DB:              21      Gaps:      0
XX
XX      JUNC_SEQ3_SEQ4 (1-30) x AAA62125 (1-201)
XX
XX      QY      1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
XX      Db      116 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCUCGACGUCC 160
XX
XX      RESULT 6
XX      AAV21727
XX      ID      AAV21727 standard; cDNA; 2125 BP.
XX
XX      AC      AAV21727;
XX
XX      DT      17-AUG-1998 (first entry)
XX
XX

```

```

DE Humanised vector pITL-hHER/neu.
XX Vector; vaccine; tumour; antigen; plasmid pITL-hHER/neu;
KW human; HER-2/neu; C-erbB-2; breast cancer; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 13..921
FT /tag= a
FT /product= human HER-2/neu
FT polyA_site 922..1181
FT /tag= b
FT /note= "Combined splice and polyA sequences"
FT CDS 1195..1401
FT /tag= c
FT /note= "SupF gene"
FT misc_feature 1412..1864
FT /tag= d
FT /note= "ColEI origin of replication"
FT promoter 1877..2125
FT /tag= e
FT /note= "RANTES promoter"
FT
XX WO9806863-A1.
XX
XX 19-FEB-1998.
XX
XX 14-AUG-1997; 97WO-US14306.
XX
XX 14-AUG-1996; 96US-0023931.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson EL, Nelson PJ;
XX WPI; 1998-159552/14.
XX
XX Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX
XX Example 7; Page 41-42; 125pp; English.
XX
XX Plasmid pITL-hHER2/neu comprises base vector pITL (see AAV211724) and
XX a human HER-2/neu nucleic sequence. pITL-hHER2/neu was used to
XX evaluate the toxicity of anti-tumour vaccination in rats, and in
XX phase I and phase II trials to evaluate polynucleotide vaccination
XX in advanced breast cancer. Novel humanised vectors, which can be
XX based on pITL, comprise a human-derived promoter or mammalian
XX homologue which is functional in mammalian target tissue and cells
XX and a sequence acceptance site which accepts cDNA products from
XX RT-PCR cloning. The vectors are non-replicating in mammalian cells
XX but are capable of extended stable expression of the target
XX sequence, generating an immune response in immunised individuals.
XX The vectors selectively elicit immune responses to the target
XX sequences with little or no immune response to the other components
XX of the vectors.
XX
XX Sequence 2125 BP; 449 A; 650 C; 598 G; 428 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.124 Length: 2125
XX Score: 80.00 Matches: 15
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 50.31% Indels: 0
XX DB: 19 Gaps: 0
XX
XX JUNC_SEQ3_SEQ04 (1-30) x AAV21727 (1-2125)
XX
XX 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

```

```

Db 112 GACAAGGGCTGCCCGCCGAGAGAGAGCCGCTGTGACGTCC 156
RESULT 7
AAZ50586
ID AAZ50586 standard; DNA; 2871 BP.
XX
XX AAZ50586;
XX
XX 23-MAY-2000 (first entry)
XX
XX DC8scFv-erbB2EC fusion construct containing tetramerisation domain.
XX
XX DC8 scFv; single-chain variable fragment; erbB2EC; extracellular domain;
XX human; fusion construct; tetramerisation domain; constant domain;
XX heteromultimer; multifunctional compound; melanoma; sarcoma;
XX immunoglobulin; cytostatic; immunostimulatory; antileukaemia; diagnosis;
XX antiproliferative; prevention; treatment; malignant; haematopoietic cell;
XX lymphoma; leukaemia; solid tumour; carcinoma; ds.
XX
XX Chimeric - Unidentified.
XX Chimeric - Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 10..2865
XX /tag= a
XX /product= "DC8scFv-erbB2EC fusion protein"
XX sig_peptide 10..66
XX /tag= b
XX /label= Leader_sequence
XX misc_feature 67..390
XX /tag= c
XX /label= DC8scFv_light_chain_variable_region
XX misc_feature 391..435
XX /tag= d
XX /label= Glycine-Serine-linker_DNA
XX misc_feature 436..771
XX /tag= e
XX /label= DC8scFv_heavy_chain_variable_region
XX misc_feature 775..807
XX /tag= f
XX /note= "5' end of human IgG3 upper hinge region
XX with 3 additional nucleotides"
XX misc_feature 808..924
XX /tag= g
XX /label= Human_p53_tetramerisation_domain
XX misc_feature 925..945
XX /tag= h
XX /label= Short_peptide_linker
XX misc_feature 946..2844
XX /tag= i
XX /label= erbB2EC_domain
XX misc_feature 2845..2862
XX /tag= j
XX /label= His_tag
XX
XX WO200006605-A2.
XX
XX 10-FEB-2000.
XX
XX 28-JUL-1999; 99WO-EP05416.
XX
XX 28-JUL-1998; 98EP-0114082.
XX (MICR-) MICROMET GES BIOMEDIZINISCHE FORSCHUNG.
XX Kufer P, Dreier T, Baeuerle PA, Borschert K, Zettl F;
XX WPI; 2000-195265/17.
XX P-PSDB; AAY44993.
XX
XX New multifunctional compounds useful for preventing and/or treating
XX malignant cell growth and for detection and diagnosis.

```

XX Example 9; Fig 49; 166pp; English.
PS The patent discloses heteroinibodies which are multifunctional compounds
XX producible in a mammalian host cell as a secretable and fully functional
CC heterodimer of two polypeptide chains, where one of the polypeptide
CC chains comprises, a CHI-domain (constant domain of an immunoglobulin
CC heavy chain) and the other chain comprises CL-domain (constant domain of
CC an immunoglobulin light chain). The polypeptide chains further comprise,
CC fused to the constant domains at least two (poly)peptides having
CC different receptor or ligand functions, where further at least two of the
CC different (poly)peptides lack an intrinsic affinity for one another and
CC are linked via the constant domains. The heteroinibodies have
CC cytostatic, immunostimulatory, antileukaemia and antiproliferative
CC activities. These compounds can be used for diagnosing, preventing and
CC treating malignant cell growth related to malignancies of haematopoietic
CC cells e.g. lymphomas and leukemias, or to solid tumours e.g. carcinomas,
CC melanomas and sarcomas.
CC The present sequence is a fusion construct comprising DC8
CC single-chain Fv (scFv) fragment at the N-terminus, extracellular
CC domain of human erbB2 at the C-terminus and a tetramerisation
CC domain between them. This construct was prepared to find out whether
CC an oligomerisation domain characterised in bacterial expression system
CC is applicable for expression of fully functional and secretable
CC recombinant protein in mammalian host cells. This tetrameric construct
CC was not expressed as secretable and fully functional protein
CC in mammalian cells. Hence general applicability of the tetramerisation
CC domain for oligomerisation strategies in mammalian cells was ruled out.
XX
SQ Sequence 2871 BP; 598 A; 868 C; 834 G; 571 T; 0 other;

Alignment Scores:
Pred. No.: 0.179 Length: 2871
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x AAZ50586 (1-2871)
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
DB 2797 GACAAGGGCTGCCCGCGCAGAGAGCCAGCCCTCTGACGTCC 2841

RESULT 8
AA89736
ID AA89736 standard; DNA; 3600 BP.
XX
AC AA89736;
XX
DT 12-JAN-2001 (first entry)
XX
DE Human HER-2/neu coding sequence.
XX
KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
KW colon cancer; ds.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 1..3600
FT /*tag= a
FT /product= "HER-2/neu protein"
XX
PN WO200044899-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-US02164.
XX
PR 29-JAN-1999; 99US-0117976.

XX (CORI-) CORIXA CORP.
PA (SMIK) SMITHLINE BEECHAM.
XX
PI Cheever MA, Gheysen D;
XX
DR WPI; 2000-505976/45.
XX P-PSDB; AAB21198, AAB21208.
XX
PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
PS Disclosure; Fig 15; 128pp; English.
XX
CC The present sequence encodes the human HER-2/neu protein. HER-2/neu is
CC a member of the tyrosine kinase family of receptor-like glycoproteins
CC and shows homology to the epidermal growth factor receptor (EGFR). It
CC probably plays a part in cell growth and/or differentiation. The
CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising
CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
XX
SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 other;

Alignment Scores:
Pred. No.: 0.235 Length: 3600
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x AA89736 (1-3600)
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
DB 1915 GACAAGGGCTGCCCGCGCAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 9
ABK86207
ID ABK86207 standard; cDNA; 3678 BP.
XX
AC ABK86207;
XX
DT 24-SEP-2002 (first entry)
XX
DE cDNA encoding human breast cancer antigen, Her2 variant.
XX
KW Human; Her2; cytostatic; antiviral; immunostimulant;
KW cell-mediated immune response; tumour; breast cancer;
KW virus infection; prostate cancer; colorectal cancer; pancreatic cancer;
KW lymphoma; leukaemia; hepatitis virus; lentivirus; herpesvirus;
KW human immunodeficiency virus; HIV; flavivirus; pestivirus; gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 7..3678
FT /*tag= a
FT /product= "Breast cancer antigen Her2 variant"
XX
PN WO200240059-A2.
XX
PD 23-MAY-2002.
XX
PF 01-NOV-2001; 2001WO-US45626.
XX
PR 01-NOV-2000; 2000US-0704232.

XX (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.
 PA (MINC/) MINCHEFF M S.
 PA (LOUK/) LOUKINOV D I.
 PA (ZOUB/) ZOUBAK S.
 XX Mincheff MS, Loukinov DI, Zoubak S;
 XX WPI: 2002-527524/56.
 DR P-PSDB; AAU98923.
 XX
 PT Inducing a cell-mediated immune response against a target antigen,
 PT reducing undesired cells and stimulating presentation of an antigen by
 PT a cell, comprises administering a polynucleotide encoding a variant of
 PT an antigen
 XX
 XX Disclosure; Page 128-134; 146pp; English.
 PS
 XX The invention relates to a method of inducing a cell-mediated immune
 CC response against a cell comprising a target antigen (I) in a subject,
 CC treating a subject having undesired cells, for example tumour cells
 CC or virally infected cells (C), reducing the number of (C) in a subject,
 CC and stimulating presentation of (I) by a cell. This is done by
 CC administering a polynucleotide (II) encoding a variant of (I), so that
 CC (II) expressed in a cell and cell-mediated immune response is induced.
 CC The method can be used to treat prostate cancer, breast cancer,
 CC colorectal cancer and pancreatic cancer, as well as lymphomas and
 CC leukaemias. The method is also useful in treating chronic viral
 CC infections such as those caused by hepatitis viruses, lentiviruses
 CC (including human immunodeficiency virus (HIV)), herpesviruses and the
 CC flaviviruses and pestiviruses. The present sequence represents the coding
 CC sequence of human breast cancer antigen, Her2 variant, used as a target
 CC antigen in the method of the invention.
 XX
 XX Sequence 3678 BP; 746 A; 1137 C; 1091 G; 704 T; 0 other;

Alignment Scores:
 Pred. No.: 0.241 Length: 3678
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 24 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x ABR86207 (1-3678)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 |||||
 Db 1825 GACAGGGGTGCCCCCGCAGCAGAGAGCCCTCTGACGTC 1869

RESULT 10
 AAT40739
 ID AAT40739 standard; cDNA; 3768 BP.
 XX
 AC AAT40739;
 XX
 XX 01-JAN-1997 (first entry)
 DT
 XX
 DE HER-2/neu oncogene.
 XX
 KW HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
 KW breast cancer; ovary cancer; colon cancer; lung cancer;
 KW prostate cancer; genetic immunisation; tumour; vaccine; vector;
 KW ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH CDS 1..3765
 FT /*tag= b
 FT /note= "nucleotides 2026-3765 (claim 1) code for
 FT HER-2/neu intracellular domain"
 XX

PN WO9630514-A1.
 XX
 PD 03-OCT-1996.
 XX
 XX 28-MAR-1996; 96WO-US01689.
 PF
 XX 31-MAR-1995; 95US-0414417.
 PR
 XX (UNIW) UNIV WASHINGTON.
 PA
 XX Cheever MA, Disis ML;
 PI
 XX WPI: 1996-455361/45.
 DR P-PSDB; AAW01111.
 DR
 XX DNA encoding HER-2-neu poly:peptide(s) - used for prevention or
 PT treatment of malignancies with which the HER-2/neu oncogene is
 PT associated
 PT
 XX Claim 1; Page 49-56; 71pp; English.
 PS
 XX Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or
 CC c-erbB2) protein (AAW01111). The oncogene is overexpressed in various
 CC cancers, including breast, ovarian, colon, lung and prostate, and
 CC appears to induce malignancies through quantitative mechanisms that
 CC result from increased or deregulated expression of an essentially
 CC normal gene product. Nucleotides 2026-3765 of the cDNA sequence
 CC code for the intracellular domain (lys676-val1255) of the HER-2/neu
 CC protein, which is useful for immunisation against malignancy.
 CC Nucleic acids can be used to direct expression of the intracellular
 CC domain in transformed host cells, or are used, alone or in a viral
 CC vector, for genetic immunisation of an animal.
 XX
 XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
 Pred. No.: 0.248 Length: 3768
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 17 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x AAT40739 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 |||||
 Db 1915 GACAGGGGTGCCCCCGCAGCAGAGAGCCCTCTGACGTC 1959

RESULT 11
 AAX01912
 ID AAX01912 standard; DNA; 3768 BP.
 XX
 AC AAX01912;
 XX
 XX 21-APR-1999 (first entry)
 DT
 XX
 DE Human HER-2/neu oncogene DNA.
 XX
 KW HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
 KW malignancy; treatment; tumour; ss.
 KW
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH CDS 1..3768
 FT /*tag= a
 FT /product= "HER-2/neu"
 FT /note= "oncogene"
 FT 2026...3765
 FT /*tag= b
 FT /note= "region which elicits immune response"
 XX

PN US5869445-A.
 XX
 PD 09-FEB-1999.
 PF
 XX 01-APR-1996; 96US-0625101.
 XX
 PR 01-APR-1996; 96US-0625101.
 PR 17-MAR-1993; 93US-0033644.
 PR 12-AUG-1993; 93US-0106112.
 PR 31-MAR-1995; 95US-0414417.
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX
 PI Cheever MA, Disis ML;
 XX
 DR WPI; 1999-152835/13.
 DR P-PSDB; AAW92406.
 XX
 PT Use of HER-2/neu polypeptides - for eliciting an immune response to
 PT an HER-2/neu associated malignancy, particularly for treating or
 PT preventing tumours
 XX
 PS Claim 1a; Column 23-32; 26pp; English.
 XX
 CC This sequence encodes the human HER-2/neu oncogene protein. A fragment
 CC of this protein is used in a method for eliciting or enhancing an immune
 CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
 CC B cells to produce an immune response to the HER-2/neu protein. The
 CC method can be used for immunisation against a malignancy in which the
 CC HER-2/neu oncogene is associated and in the treatment of an existing
 CC tumour, or to prevent tumour occurrence or reoccurrence.
 XX
 SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.248 Length: 3768
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 20 Gaps: 0
 JUNC_SEQ3_SEQ4 (1-30) x AAX01912 (1-3768)
 QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 |||||
 DB 1915 GACRAGGGCTGCCCGCGAGCAGAGACCGACCCCTCTGACGTCC 1959
 RESULT 12
 AAA09455
 ID AAA09455 standard; DNA; 3768 BP.
 XX
 AC AAA09455;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Human heregulin 2 (Her2) coding sequence.
 XX
 KW Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
 KW self-protein; cell-associated peptide antigen; foreign epitope;
 KW cancer; breast cancer; prostate cancer; ss.
 OS Homo sapiens.
 XX
 PN WO200020027-A2.
 XX
 PD 13-APR-2000.
 XX
 PF 05-OCT-1999; 99WO-DK00525.
 XX
 PR 05-OCT-1998; 98DK-0001261.
 PR 20-OCT-1998; 98US-0105011.
 XX

PA (MEBI-) M & E BIOTECH AS.
 XX
 PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX
 DR WPI; 2000-349917/30.
 DR P-PSDB; AAY92620.
 XX
 PT Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 XX
 PS Claim 62; Page 187-193; 220pp; English.
 XX
 CC The claims detail a method for inducing immune responses against weakly
 CC immunogenic cell-associated peptide antigens (PA) such as those
 CC associated with cancers (i.e. self-proteins), for example, human
 CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
 CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
 CC simultaneous presentation by antigen producing cells (APCs) of the
 CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
 CC group derived from the PA and/or at least 1 B-cell group derived from the
 CC cell-associated PA; and (2) at least 1 first T helper cell group which is
 CC foreign to the animal. Analogues of human PSM, human Her2 and
 CC human/murine FGF8b comprising a substantial part of all known and
 CC predicted CTL and B-cell epitopes of the respective PA and including at
 CC least one foreign T helper epitope are also claimed. The method is used
 CC to treat prostate, prostate/breast or breast cancer when the PA is human
 CC PSM, FGF8b and Her2, respectively.
 XX
 SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.248 Length: 3768
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 21 Gaps: 0
 JUNC_SEQ3_SEQ4 (1-30) x AAA09455 (1-3768)
 QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 |||||
 DB 1915 GACRAGGGCTGCCCGCGAGCAGAGACCGACCCCTCTGACGTCC 1959
 RESULT 13
 AAH23392
 ID AAH23392 standard; DNA; 3768 BP.
 XX
 AC AAH23392;
 XX
 DT 25-SEP-2001 (first entry)
 XX
 DE Human HER-2/neu protein encoding DNA.
 XX
 KW Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
 KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3768
 FT /*tag= a
 FT /product= "HER-2/neu protein"
 XX
 PN WO200153463-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 19-JAN-2001; 2001WO-US01850.
 PR 21-JAN-2000; 2000US-0177545.
 XX

PA (CORI-) CORIXA CORP.
 XX
 PI Cheever MA, Hand-Zimmermann S;
 XX
 DR WPI: 2001-476112/51.
 DR P-PSDB: AAB83458.
 XX
 PT New antigen-presenting cells, useful as vaccines for eliciting or
 PT enhancing an immune response to HER-2/neu protein, particularly useful
 PT for treating or preventing cancer, e.g. breast cancer -
 XX
 PS Claim 1; Page 41-46; 49pp; English.
 XX
 CC The invention provides an isolated antigen-presenting cell, which
 CC expresses at least an immunogenic portion of a polypeptide that produces
 CC an immune response to HER-2/neu protein. The antigen-presenting cells are
 CC useful as vaccines for eliciting or enhancing an immune response to
 CC HER-2/neu protein, particularly in treating or preventing malignancies in
 CC which the HER-2/neu oncogene is associated. Specifically, these are
 CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,
 CC colon, lung or prostate cancers. The present sequence represents a DNA
 CC encoding the human HER-2/neu protein (also known as p185 or c-erbB2).
 XX
 SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.248 Length: 3768
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 22 Gaps: 0
 JUNC_SEQ3_SEQ4 (1-30) x AAB23392 (1-3768)
 QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 ID AB235744 standard; DNA; 3768 BP.
 XX
 AC AB235744;
 XX
 DT 07-FEB-2003 (first entry)
 XX
 DE Human ERBB2 polynucleotide SEQ ID NO 52.
 XX
 KW Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
 KW protozoacide; gene expression; antisense; tumour; infection; Plasmodium;
 KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
 KW Hepatitis C virus; human papilloma virus; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN DE10100588-A1.
 XX
 PD 18-JUL-2002.
 XX
 PF 09-JAN-2001; 2001DE-1000588.
 XX
 PR 09-JAN-2001; 2001DE-1000588.
 XX
 PA (RIBO-) RIBOPHARMA AG.
 XX
 PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
 XX
 DR WPI: 2002-683450/74.
 XX
 PT Inhibiting expression of target genes, useful e.g. for treating tumors,
 PT by introducing into cells two double-stranded RNAs that are
 PT complementary to the target -

XX Claim 13; Page 38-39; 100pp; German.
 PS
 CC The invention relates to inhibiting expression of a target gene in a cell
 CC by introducing at least two oligoribonucleotides (dsRNAi and II), both
 CC with a double-stranded (ds) structure of at most 49 sequential nucleotide
 CC pairs. At least part of one strand (S1, S2) of the ds structures in each
 CC of dsRNAi and II are complementary to regions in the target gene. The
 CC method uses antisense inhibition of gene expression using double stranded
 CC RNA inhibition (RNAi). The method is particularly used to treat tumours
 CC or infections, especially by Plasmodium or viruses/viroids (pathogenic on
 CC humans, animals or plants). The method provides more effective inhibition
 CC of expression than known methods using a single dsRNA, even at very low
 CC concentrations. When dsRNA has at least one unpaired nucleotide at the
 CC end, stability (and thus effective concentration in the cell) is
 CC improved and efficiency can be increased further by pretreating the cells
 CC with interferon. The present sequence is that of a target DNA of the
 CC invention.
 XX
 SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.248 Length: 3768
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.31% Indels: 0
 DB: 24 Gaps: 0
 JUNC_SEQ3_SEQ4 (1-30) x AB235744 (1-3768)
 QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 ID ABX09987
 DB 1915 GACAAGGGCTGCCCGCGGACGAGAGCCAGCCCTCTGACGTC 1959
 RESULT 15
 ABX09987
 ID ABX09987 standard; DNA; 3768 BP.
 XX
 AC ABX09987;
 XX
 DT 23-JAN-2003 (first entry)
 XX
 DE Human ERBB2 DNA fragment SEQ ID 52.
 XX
 KW Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;
 KW prion; inhibition; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN DE10100587-Cl.
 XX
 PD 21-NOV-2002.
 XX
 PF 09-JAN-2001; 2001DE-1000587.
 XX
 PR 09-JAN-2001; 2001DE-1000587.
 XX
 PA (RIBO-) RIBOPHARMA AG.
 XX
 PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
 XX
 DR WPI: 2002-742209/81.
 XX
 PT Inhibiting expression of target genes, e.g. oncogenes, in cells, by
 PT introduction of complementary double-stranded oligoribonucleotide,
 PT after treating the cell with interferon -
 XX
 PS Disclosure; Page 43-44; 98pp; German.
 XX
 CC This invention describes a novel method for inhibiting expression of a
 CC target gene by introducing into the cell that contains the target gene
 CC at least one oligoribonucleotide (dsRNAi) that has a double-stranded

CC (ds) structure of not more than 49 consecutive nucleotides (nt), where
CC at least a segment of one strand of the ds structure is complementary
CC with the target gene and the cells are treated with interferon before
CC introduction of dsRNA. The method is used to inhibit expression of
CC target genes, particularly oncogenes, cytokine genes, id (not defined)
CC protein genes; developmental or prion genes, or genes expressed in
CC pathogenic organisms (particularly plasmidia) or in viruses or viroids
CC (pathogenic in humans, animals or plants). Treating the cells with
CC interferon greatly increases the extent to which dsRNA can inhibit
CC expression of the target genes, and the effect is even greater when dsRNA
CC are modified to increase their stability. ABX0936-ABX10075 represent
CC gene fragments used to illustrate the method of the invention.
xx

SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:			
Pred. No.:	0.248	Length:	3768
Score:	80.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.31%	Indels:	0
DB:	24	Gaps:	0
JUNC_SEQ3_SEQ4 (1-30) x ABX09987 (1-3768)			
QY	1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15		
Db	 1915 GACAAGGGCTGCCCCCGCGAGCAGAGAGCCCTCTGACGTCC 1959		

Search completed: October 15, 2003, 23:48:38
Job time : 31.4941 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 432.026 Seconds
(without alignments)
2840.777 Million cell updates/sec

Title: JUNC_SEQ3_SEQ4

Perfect score: 159

Sequence: 1 DKGCPAEQRASPLTSQEDLGPA SPLDSTF 30

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/HOLLERAN480/runat_15102003_131912_20536/app_query.fasta_1.4685
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 @CGN_1_1_22724_runat_15102003_131912_20536 -NCPU=6 -ICPU=3
-NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_ov.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_or.*

21: em_ov.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	144	90.6	2763	6	AX380942	AX380942 Sequence
2	144	90.6	2781	6	AX380944	AX380944 Sequence
C 3	85.5	53.8	186134	10	AL591390	AL591390 Mouse DNA
C 4	85.5	53.8	214019	2	AC064803	AC064803 Mus muscu
5	85	53.5	13450	9	AB096612	AB096612 Homo sapi
6	85	53.5	16572	9	AB096614	AB096614 Homo sapi
7	85	53.5	20271	9	AB096613	AB096613 Homo sapi
8	85	53.5	30837	9	AY208911	AY208911 Homo sapi
9	85	53.5	161815	9	AC079199	AC079199 Homo sapi
10	85	53.5	168585	9	AC040933	AC040933 Homo sapi
11	85	53.5	198008	2	AC142197	AC142197 Homo sapi
12	84	52.8	2385	6	AR082744	AR082744 Sequence
13	84	52.8	2385	6	AR099963	AR099963 Sequence
14	84	52.8	2385	6	AR143949	AR143949 Sequence
15	80	50.3	153	6	AX022984	AX022984 Sequence
16	80	50.3	201	6	AR229723	AR229723 Sequence
17	80	50.3	281	6	AX023363	AX023363 Sequence
18	80	50.3	3678	6	AX051114	AX051114 Sequence
19	80	50.3	3768	6	AR034479	AR034479 Sequence
20	80	50.3	3768	6	AX060704	AX060704 Sequence
21	80	50.3	3768	6	AX201817	AX201817 Sequence
22	80	50.3	3768	6	AX380923	AX380923 Sequence
23	80	50.3	3768	6	AX384604	AX384604 Sequence
24	80	50.3	3768	6	AX465456	AX465456 Sequence
25	80	50.3	3768	6	AX467229	AX467229 Sequence
26	80	50.3	3768	6	AX481438	AX481438 Sequence
27	80	50.3	4473	6	AR080259	AR080259 Sequence
28	80	50.3	4473	6	AR167390	AR167390 Sequence
29	80	50.3	4473	9	HSEB2R	X03363 Human c-erb
30	80	50.3	4530	6	AR202597	AR202597 Sequence
31	80	50.3	4530	6	AR283481	AR283481 Sequence
32	80	50.3	4530	6	AX282577	AX282577 Sequence
33	80	50.3	4530	6	AX587649	AX587649 Sequence
34	80	50.3	4530	6	AX644071	AX644071 Sequence
35	80	50.3	4530	6	BD005474	BD005474 Cellular
36	80	50.3	4530	6	I21124	I21124 Sequence 9
37	80	50.3	4530	6	I59745	I59745 Sequence 9
38	80	50.3	4530	9	HUMHER2A	M11730 Human tyros
39	80	50.3	9274	6	AX060703	AX060703 Sequence
40	79	49.7	1755	6	AX384609	AX384609 Sequence
41	79	49.7	1757	6	AX384607	AX384607 Sequence
42	79	49.7	1773	6	AX384610	AX384610 Sequence
43	79	49.7	1806	6	AX384608	AX384608 Sequence
44	79	49.7	3780	4	AB008451	AB008451 Canis fam
45	76	47.8	4062	10	HANNEU	D16295 Mesocricetu

ALIGNMENTS

RESULT 1

```
JUNC_SEQ3_SEQ4 (1-30) x X380944 (1-2781)
```

DY
DZ

1 AsplyGlyCysProAlaGluInArGrAlAserProLeuThrSerGlnAnSnGIuAaspLeu 20
:::~::~:
1918 GAACAGGGTCCCCAGCACAGACAGAGACCAGCCCGCTGTCTCAGACGAGACTTA 1977

DY
DZ

21 GlyProAlaSerProLeuAaspSserThrPhe 30
|||||:::~::::||
1978 GGCCCCGCCAGCCCATTGGACAGCACTTC 2007

RESULT 3
AL591390/c
LOCUS
DEFINITION Mouse DNA sequence from clone RP23-355L10 on chromosome 11,
complete sequence.
ACCESSION AL591390
VERSION AL591390.8 GI:17017767
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
Phillimore,B.
Direct Submission
Submitted (17-Nov-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquies@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 20, 2001 this sequence version replaced gi:1655512.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name(s). Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WormPEP; Information on the WormPEP
database can be found at
<http://www.sanger.ac.uk/projects/C-elegans/wormpep/> RP23-355L10 is
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6

IMPORTANT! This sequence is not the entire insert of clone
RP23-355L10 it may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP23-355L10 is at 186134 in this
sequence. The true right end of clone RP23-438D7 is at 2000 in this
sequence..

FEATURES
source Location/Qualifiers
1 .186134
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-355L10"
/clone_lib="RPCI-23"
17991 ..18044
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."
misc_feature 47794..47812
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."

BASE COUNT 46287 a 47900 c 47689 g 44258 t
ORIGIN

gene	/note="3', exon"	
exon	11349..11517	
	/gene="GRB7"	
	11349..11517	
	/gene="GRB7"	
BASE COUNT	2748 a 3818 c 3387 g 3496 t	1 others
ORIGIN		
Alignment Scores:		
Pred. No.:	0.415	Length: 13450
Score:	85.00	Matches: 23
Percent Similarity:	46.00%	Conservative: 0
Best Local Similarity:	46.00%	Mismatches: 3
Query Match:	53.46%	Indels: 24
DB:	9	Gaps: 2
JUNC_SEQ3_SEQ4 (1-30) x AB0966612 (1-13450)		
Qy	5 ProAlaGlu	-----GlnArgAlaSerProLeu
Db	96 CCTGCAGAGGGTGGGAAGGAGAGATGATCAGTATCCAGTATCCAGGCCGCCCTC	
Qy	16 -----	-----Gln
Db	156 ATGCTGGGCTGGGGAGGGGCCACCATCTGCCTCTCCTTCCTCCTCCACAG	
Qy	21 GlyProAlaSerProLeuAspSerThrPhe 30	
Db	216 GGCCCGACCCAGTCCCTTGGAGACACCTTC 245	
RESULT 6		
AB0966614	AB0966614 16572 bp DNA linear	
LOCUS	Homo sapiens DNA, 16.5kb-T5 rearranged sequence.	
DEFINITION	AB0966614	
ACCESSION	AB0966614.1 GI:30962527	
VERSION	Homo sapiens (human)	
KEYWORDS	Homo sapiens	
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homninda	
ORGANISM	1	
REFERENCE	Kawahara.Y., Tanabe.C., Ikeuchi.T., Aoyagi.K., N. Sakamoto.H., Hoshinaga.K., Yoshida.T., Sasaki.H. et al. Presence of novel mechanisms other than the BFB cy amplification of human oncogene unpublished	
AUTHORS	2 (bases 1 to 16572)	
TITLE	Sasaki,H.	
JOURNAL	Direct Submission	
REFERENCE	Submitted (20-NOV-2002) Hiroki Sasaki, National Cancer Research Institute, Genetics Division; Tsukiji 5-1 Tokyo 104-0045, Japan (E-mail: hksasaki@gan2.res.nco. Tel:81-3-3542-2511(ex.4402), Fax:81-3-3541-2685)	
AUTHORS	Location/Qualifiers	
TITLE	1. 16572	
JOURNAL	/organism="Homo sapiens"	
REFERENCE	/mol_type="genomic DNA"	
AUTHORS	/db_xref="taxon:9606"	
TITLE	/chromosome="17"	
JOURNAL	/map="17q12"	
FEATURES	/clone="16.5kb-T5 rearranged DNA"	
source	684..2046	
gene	/gene="c-ERBB-2"	
exon	684..936	
exon	1078..2046	
gene	/gene="c-ERBB-2"	
exon	/note="3', exon"	
exon	11349..11517	
exon	/gene="GRB7"	
exon	11349..11517	

misc_feature		11419		/gene="GRB7"		/note="inter-amplicon junction"	
gene		14639..14711		/gene="c-ERBB-2"			
exon		14639..14711		/gene="c-ERBB-2"			
BASE COUNT		4372 a 5349 c 5309 g 5241 t					
ORIGIN							
Alignment Scores:							
Pred. No.:		0.629		Length:		20271	
Score:		85.00		Matches:		23	
Percent Similarity:		46.00%		Conservative:		0	
Best Local Similarity:		46.00%		Mismatches:		3	
Query Match:		53.46%		Indels:		24	
DB:		9		Gaps:		2	
JUNC_SEQ3_SEQ4 (1-30) x AB096613 (1-20271)							
QY		5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer-----		15			
Db		96 CTGCGAGAGGGTGGGAGGAGAGATGAGTCAGTATGCCAGGCCCCCTCACGGAAGGCTGC		155			
QY		16 -----GlnAsnGluAspLeu 20					
Db		156 ATGCTGGGCTGGGAGGGCCACCATCTGCTCTCTCTCCACAGAAATGAGGACTTG		215			
RESULT 8							
AY208911		21 GlyProAlaSerProLeuAspSerThrPhe 30					
LOCUS		Homo sapiens v-erb-b2 erythroblastic leukemia viral oncogene					
DEFINITION		homolog 2, neuro/glioblastoma derived oncogene homolog (avian					
ACCESSION		AY208911					
VERSION		AY208911.1					
KEYWORDS		Homo sapiens (human)					
SOURCE		Homo sapiens					
ORGANISM		Homo sapiens					
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE		1 (bases 1 to 30837)					
JOURNAL		Rieder,M.J., Livingston,R.J., Daniels,M.R., Montoya,M.A.,					
FEATURES		Chung,M.-W., Miyamoto,K.E., Nguyen,C.P., Nguyen,D.A., Poel,C.L.,					
source		Robertson,P.D., Schackwitz,W.S., Sherwood,J.K., Witrak,L.A. and					
1..30837		Nickerson,D.A.					
/organism="Homo sapiens"		Direct Submission					
/mol_type="genomic DNA"		Submitted (30-DEC-2002) Genome Sciences, University of Washington,					
/db_xref="taxon:9606"		1705 NE Pacific, Seattle, WA 98195, USA					
/map="17q12"		To cite this work please use: NIEHS-SNPs, Environmental Genome					
/clone="20kb-MKN7 rearranged DNA."		Project, NIEHS ES15478, Department of Genome Sciences, Seattle, WA					
/cell_line="MKN7"		(URL: http://egp.gs.washington.edu).					
/tissue_type="gastric cancer"							
684..2046							
/gene="c-ERBB-2"							
684..936							
/gene="c-ERBB-2"							
1078..2046							
/gene="c-ERBB-2"							
/note="3' exon"							
11349..11419							
/gene="GRB7"							
11349..11419							
/gene="GRB7"							


```

/replace="g"
/702..9770
/rpt_family="tip100"
/rpt_type=dispersed
/766
/gene="ERBB2"
/frequency="0.01"
/replace="t"
/799..9946
/rpt_family="tip100"
/rpt_type=dispersed
/949..10151
/rpt_family="MIR"
/rpt_type=dispersed
/10009
/gene="ERBB2"

```

```

Alignment Scores:
Pred. No.: 0.963 Length: 30837
Score: 85.00 Matches: 23
Percent Similarity: 46.00% Conservative: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match: 53.46% Indels: 24
Db: 9 Gaps: 2

```

JUNC_SEQ3_SEQ4 (1-30) x AY208911 (1-30837)

```

QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
|||||
Db 27839 CCTGAGAGGTGGGAGGAGATGATCCATGATGCCAGCCCTCACGAGGCGTC 27898
QY 16 -----GlnAsnGluAspLeu 20
|||||
Db 27899 ATGTGGCTGGGGGGGCCACCATCTGCTCTCTCTCTCCACAGAATGAGGACTTG 27958
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db 27959 GCGCCAGCCAGTCCCTTGGACAGCACCTTC 27988

```

```

RESULT 9
AC079199 AC079199 161815 bp DNA linear PRI 02-OCT-2002
LOCUS Homo sapiens chromosome 17, clone RP11-94L15, complete sequence.
DEFINITION AC079199
ACCESSION AC079199
VERSION AC079199.9 GI:23462913
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 161815)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-94L15
Unpublished
2 (bases 1 to 161815)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bida,F., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,R., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Rieback,M., Riley,R.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Severy,P.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,

```

TITLE
JOURNAL
REFERENCE
AUTHORS

Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-AUG-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 161815)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission

Submitted (08-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 161815)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission

Submitted (02-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 2, 2002 this sequence version replaced gi:22758800.

All repeats were identified using RepeatMasker:

Snit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: LI0635

Center clone name: 94_L_15

FEATURES
source

Location/Qualifiers
1. 161815
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-94L15"
/clone_l1b="RPC1-11 Human Male BAC"
7. .470
/note="PCR product sequence only"

unsure


```

repeat_region complement(262..317)
unsure      /rpt_family="MIR"
repeat_region /note="PCR product sequence only"
repeat_region 1689..1986
repeat_region /rpt_family="AluX"
repeat_region 5033..5179
repeat_region /rpt_family="MIR"
repeat_region 6113..6233
repeat_region /rpt_family="MIR"
repeat_region 6418..6460
repeat_region /rpt_family="L3"
repeat_region 7910..7930
repeat_region /rpt_family="AT-rich"
repeat_region 8422..8476
repeat_region /rpt_family="(CCTG)n"
repeat_region 8513..8659
repeat_region /rpt_family="MIR"
repeat_region complement(8917..9066)
repeat_region /rpt_family="MIR3"
repeat_region 12588..12697
repeat_region /rpt_family="G-rich"
repeat_region complement(13513..13824)
repeat_region /rpt_family="AluSp"
repeat_region complement(13836..13937)
repeat_region /rpt_family="MIR"
repeat_region 18739..18939
repeat_region /rpt_family="MIR"
repeat_region 21075..21274
repeat_region /rpt_family="MIR"
repeat_region 21447..21714
repeat_region /rpt_family="MIR"
repeat_region complement(21385..23684)
repeat_region /rpt_family="AluX"
repeat_region 24620..24819
repeat_region /rpt_family="MIR"
repeat_region 24991..25298
repeat_region /rpt_family="L2"
repeat_region complement(25299..25605)
repeat_region /rpt_family="AluJb"
repeat_region 25606..25686
repeat_region /rpt_family="L2"
repeat_region 25687..25838
repeat_region /rpt_family="MIR"
repeat_region 25839..25952
repeat_region /rpt_family="LTR54B"
repeat_region complement(25953..26241)
repeat_region /rpt_family="AluSp"
repeat_region 26242..26378
repeat_region /rpt_family="LTR54B"
repeat_region 26379..26695
repeat_region /rpt_family="AluY"
repeat_region 26696..26901
repeat_region /rpt_family="LTR54B"
repeat_region 26902..26985
repeat_region /rpt_family="MIR"
repeat_region 26986..27224
repeat_region /rpt_family="L2"
repeat_region 27243..27550
repeat_region /rpt_family="AluSg"
repeat_region 27695..28007
repeat_region /rpt_family="L2"
repeat_region 28290..28600
repeat_region /rpt_family="AluX"
repeat_region 29462..29600
repeat_region /rpt_family="MIR"
repeat_region 29643..29967
repeat_region /rpt_family="AluY"
repeat_region complement(30700..31000)
repeat_region /rpt_family="AluSp"
repeat_region complement(31002..31248)
repeat_region /rpt_family="AluJb"
repeat_region 31521..31572

```

```

repeat_region /rpt_family="L2"
31577..31796
repeat_region /rpt_family="MIR"
32751..33072
repeat_region /rpt_family="L2"
33181..33293
repeat_region /rpt_family="L2"
34072..34376
repeat_region /rpt_family="AluSp"
complement(34542..34833)
repeat_region /rpt_family="AluJb"
complement(35218..35297)
repeat_region /rpt_family="L2"
35320..35431
repeat_region /rpt_family="AluSp"
35433..35500
repeat_region /rpt_family="AluSp/q"
complement(35581..35702)
repeat_region /rpt_family="L2"
35710..36008
repeat_region /rpt_family="AluSc"
36009..36023

Alignment Scores:
Pred. No.:      5.18      Length:      161815
Score:          85.00     Matches:      23
Percent Similarity: 46.00%  Conservative: 0
Best Local Similarity: 46.00%  Mismatches:   3
Query Match:      53.46%    Indels:      24
DB:               9        Gaps:        2

JUNC_SEQ3_SEQ4 (1-30) x AC079199 (1-161815)
QY      5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
       |||||
Db      71166 CCTGCAGAGGGTGGGAAGGAGAGATGATCCAGTATGCCAGGCCCTCAGGAAGCTGC 71225
QY      16 -----GlnAsnGluAspLeu 20
       |||||
Db      71226 ATGCTGGGCTGGGGAGGGCCACCATCTCTCTCTCCCTCCACAGATGAGGACTTG 71285
QY      21 GlyProAlaSerProLeuAspSerThrPhe 30
       |||||
Db      71286 GCGCCAGCCAGTCCCTTGGACACGACCTTC 71315

RESULT 10
AC040933
LOCUS      Homo sapiens chromosome 17, clone CTD-2019C10, complete sequence.
DEFINITION AC040933
ACCESSION AC040933
VERSION   AC040933.12 GI:23396303
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 168585)
AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
TITLE     Homo sapiens chromosome 17, clone CTD-2019C10
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 168585)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
           Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
           Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
           Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
           Collumore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
           Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
           Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
           Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
           Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
           Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
           Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
           McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,

```


[illegible]

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db 10 GACAAGGGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGAGGTCC 54

Search completed: October 16, 2003, 10:59:54
Job time : 503.208 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 234.794 Seconds
(without alignments)
3105.423 Million cell updates/sec

Title: JUNC_SEQ3_SEQ2RES991_

Perfect score: 160

Sequence: 1 DRGCPAEQRAPLTSQNDLGPSPMDSTF 30

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.p2n.model -DRV=rlp
-Q/cgn2_1/USPTO.spool_p/HOLLERAN480/runat_15102003_131913_20548/app_query.fasta_1.4685
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 @CGN_1_1_12645 @runat_15102003_131913_20548 -NCPU=6 -ICPU=3
-NO_MMALP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	91.5	57.2	618	28	AZ699313	AZ699313 RPCI-23-2
2	83.5	52.2	673	12	BI157930	BI157930 602920909
3	80	50.0	378	10	BF869854	BF869854 IL3-ET011
4	80	50.0	411	2	HS069572	Bx479682 Homo sapi
5	80	50.0	422	9	AW606421	AW606421 RC3-HT037
6	80	50.0	422	10	BF554233	BF554233 UI-R-C0-h
7	80	50.0	433	10	BF762317	BF762317 IL2-CS004
8	80	50.0	439	10	BE157169	BE157169 RC3-HT037
9	80	50.0	450	12	BG988031	BG988031 MR2-HT116
10	80	50.0	466	12	BM720186	BM720186 UI-E-E00-
11	80	50.0	524	13	BQ828479	BQ828479 LL6in1002
12	80	50.0	527	14	CB719751	CB719751 AMGNNUC:N
13	80	50.0	546	12	BI194790	BI194790 602948150
14	80	50.0	547	9	AA443351	AA443351 zw85b05.r
15	80	50.0	561	12	BM720098	BM720098 UI-E-E00-
16	80	50.0	568	13	BQ831943	BQ831943 LL6in1122
17	80	50.0	577	13	BQ829915	BQ829915 LL6in2239
18	80	50.0	592	10	BE382065	BE382065 601272225
19	80	50.0	607	12	BM083444	BM083444 Imageqc_6
20	80	50.0	613	9	AW410534	AW410534 fh06h06.x
21	80	50.0	629	10	BE374596	BE374596 601225384
22	80	50.0	666	13	BU671195	BU671195 NISC_lr06
23	80	50.0	666	13	BU671195	BU671195 NISC_lr06
24	80	50.0	667	12	BM048624	BM048624 603623984
25	80	50.0	674	12	BM758528	BM758528 K-EST0038
26	80	50.0	715	13	BU052728	BU052728 UI-M-EXO-
27	80	50.0	726	12	BI648082	BI648082 603279290
28	80	50.0	741	10	BE615590	BE615590 601278973
29	80	50.0	770	12	BI645946	BI645946 603274921
30	80	50.0	773	12	BI157032	BI157032 602921306
31	80	50.0	778	12	BG863431	BG863431 60297093
32	80	50.0	785	12	BI688650	BI688650 603311262
33	80	50.0	787	12	BI156730	BI156730 602922150
34	80	50.0	844	12	BI158626	BI158626 602922395
35	80	50.0	881	9	AI906106	AI906106 RC-BT105-
36	80	50.0	881	12	BI645770	BI645770 603275203
37	80	50.0	886	13	BQ882315	BQ882315 AGENCOURT
38	80	50.0	887	14	CA980253	CA980253 AGENCOURT
39	80	50.0	893	14	CA487631	CA487631 AGENCOURT
40	80	50.0	905	13	BU541491	BU541491 AGENCOURT
41	80	50.0	909	10	BG289457	BG289457 602381495
42	80	50.0	922	14	CA453990	CA453990 AGENCOURT
43	80	50.0	944	14	CD558559	CD558559 AGENCOURT
44	80	50.0	984	13	BU523863	BU523863 AGENCOURT
45	80	50.0	1009	13	BQ647343	BQ647343 AGENCOURT

ALIGNMENTS

RESULT 1
AZ699313 618 bp DNA linear GSS 24-JAN-2001
LOCUS RPCI-23-240N16.TJ RPCI-23 Mus musculus genomic clone RPCI-23-240N16
DEFINITION , genomic survey sequence.
ACCESSION AZ699313
VERSION AZ699313.1 GI:12419623
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 618)

AUTHORS Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akintret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P., and Fraser, C. M.

TITLE Mouse BAC End Sequences from Library RPCI-23

JOURNAL Unpublished

COMMENT Other_GSSS: RPCI-23-240N16.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 240 row: N column: 16
Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers
1..618
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-240N16"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site.1: EcoRI; Site.2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methyase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 110 a 214 c 151 g 143 t
ORIGIN

Alignment Scores:

Pred. No.: 0.057 Length: 618
Score: 91.50 Matches: 20
Percent Similarity: 72.41% Conservative: 1
Best Local Similarity: 68.97% Mismatches: 5
Query Match: 57.19% Indels: 3
DB: 28 Gaps: 1

JUNC_SEQ3_SEQ2RES991_ (1-30) x AZ699313 (1-618)

QY 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21
Db 165 GAGGCGTGT-----CACCTCGTGTCTCCCTTCTCTCTAGAACGAGGACTTAGGC 215
QY 22 ProSerProMetAspSerThrPhe 30
Db 216 CCTCCAGCCCATGACGACGACCTTC 242

RESULT 2

BI157930
LOCUS BI157930 673 bp mRNA linear EST 05-JUL-2001
DEFINITION 602920909F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5061027 5', mRNA sequence.
ACCESSION BI157930
VERSION BI157930.1 GI:14617931
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 673)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

JOURNAL Unpublished
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM1165 row: b column: 04
High quality sequence stop: 664.

FEATURES

Location/Qualifiers
1..673
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5061027"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT5; Site.1: SalI; Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT 146 a 198 c 181 g 148 t
ORIGIN

Alignment Scores:
Pred. No.: 0.668 Length: 673
Score: 83.50 Matches: 20
Percent Similarity: 73.33% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 4
Query Match: 52.19% Indels: 4
DB: 12 Gaps: 1

JUNC_SEQ3_SEQ2RES991_ (1-30) x BI157930 (1-673)

QY 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21
Db 473 GAGGCGTGT-----CACCTCGTGTCTCCCTTCTCTCTAGAACGAGGACTTAGG 523
QY 21 yProSerProMetAspSerThrPhe 30
Db 524 CCTCCAGCCCATGACGACGACCTTC 551

RESULT 3

BF869854/c
LOCUS BF869854 378 bp mRNA linear EST 17-JAN-2001
DEFINITION IL3-ET0114-251000-317-B07 ET0114 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF869854
VERSION BF869854.1 GI:12259984
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 378)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsumoto, A., Baia, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663


```

/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0371"
/Note="Organ: head,neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      77 a 126 c 135 g 84 t
ORIGIN

Alignment Scores:
Pred. No.:      1.04      Length:      422
Score:          80.00     Matches:      21
Percent Similarity: 46.00%   Conservatives: 2
Best Local Similarity: 42.00% Mismatches: 3
Query Match:      9      Indels:      24
DB:               9      Gaps:        2

JUNC_SEQ3_SEQ2RES991_ (1-30) x AW606421 (1-422)

QY      5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
|||||
Db      256 CCTGCAGAGGGTGGGAAGGAGAGATGAGTCCAGATGCGCAGGCCCTCAGCGAAGGCTGC 315
|||||

QY      16 -----GlnAsnGluAspLeu 20
|||||

Db      316 ATGCTGGGTGGGAGGGGCCACCATCTGCTCTCTCTCCATCCACAGATGAGGACTTG 375
|||||

QY      21 GlyProSerProMetAspSerThrPhe 30
|||||

Db      376 GCGCCAGCAGTCCCTTGGACAGCACCTTC 405
|||||

RESULT 6
BF554233
LOCUS      422 bp mRNA linear EST 12-DFC-2000
DEFINITION UI-R-C0-hp-c-01-0-UI.r1 UI-R-C0 Rattus norvegicus cDNA clone
ACCESSION BF554233
VERSION BF554233.1 GI:11663963
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 422)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LNL (infoimage.lnl.gov). IMAGE ID- 1781731
Seq primer: M13 Forward.
Location/Qualifiers
1. .422
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"

```

```

/clone="UI-R-C0-hp-c-01-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-C0"
/Note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C0
library is a subtracted library derived from the UI-R-A1
and UI-R-E1 libraries. The UI-R-A1 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, and muscle. The UI-R-E1
library consisted of a mixture of individually tagged
normalized libraries constructed from 8, 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-dT track which
allows identification of the library of origin of a clone
within the mixture. The subtracted library (UI-R-C0) was
constructed as follows: PCR amplified cDNA inserts from a
pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had
been derived was used as a driver in a hybridization with
the pooled UI-R-A1 and UI-R-E1 library in the form of
single-stranded circles. The remaining single-stranded
circles (subtracted library) was purified by
hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the UI-R-C0
library. This procedure has been previously described
(Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
1996)."
BASE COUNT      90 a 115 c 125 g 92 t
ORIGIN

Alignment Scores:
Pred. No.:      1.04      Length:      422
Score:          80.00     Matches:      15
Percent Similarity: 100.00%   Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:      10      Indels:      0
DB:               10      Gaps:        0

JUNC_SEQ3_SEQ2RES991_ (1-30) x BF554233 (1-422)

QY      16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
|||||

Db      131 CAGACGAGGACTTGGCCCATCCAGCCCATGGACAGTACCTTC 175
|||||

RESULT 7
BF762317/c
LOCUS      433 bp mRNA linear EST 12-JAN-2001
DEFINITION IL2-CS0049-181000-202-A06 CS0049 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF762317
VERSION BF762317.1 GI:12110217
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 433)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

```

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL2&tl2=IL2-CS0049-181000-202-A06&t3=2000-02-25&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 387.

FEATURES

Location/Qualifiers
 1..433
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="CS0049"

/note="Organ: colon_est; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 86 a 149 c 108 g 90 t

ORIGIN

Alignment Scores:
 Pred. No.: 1 08 Length: 433
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.00% Indels: 0
 DB: 10 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x BF762317 (1-433)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 367 GACAAGGCTGCGCCCGCAGCAGACGACGCCCTCTGACGTC 323

RESULT 8

BE157169/c

LOCUS RC3-HT0371-250200-016-a11 HT0371 Homo sapiens cDNA, mRNA linear EST 21-JUN-2000

DEFINITION BE157169

ACCESSION BE157169

VERSION BE157169.1 GI:8619890

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 439)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC3-HT0371-250
 200-016-a11&t3=2000-02-25&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 439.

FEATURES

Location/Qualifiers
 1..439
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HT0371"

/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 89 a 139 c 132 g 79 t

ORIGIN

Alignment Scores:
 Pred. No.: 1 09 Length: 439
 Score: 80.00 Matches: 21
 Percent Similarity: 46.00% Conservative: 2
 Best Local Similarity: 42.00% Mismatches: 3
 Query Match: 50.00% Indels: 24
 DB: 10 Gaps: 2

JUNC_SEQ3_SEQ2RES991_ (1-30) x BE157169 (1-439)

QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15

Db 167 CCTGCAGAGGCTGGGAAGGAGAGATGATCCAGTATGCCAGGCCCTCACGAGACTTG 48

QY 16 -----GlnAsnGluAspLeu 20

Db 107 ATGCTGGGCTGGGAGGGCCACCATCTCTCTCTCTCCATCCACAGATGAGGACTTG 48

QY 21 GlyProSerSerProMetAspSerThrPhe 30

Db 47 GCGCCAGCAGTCCCTTGGACAGCACCTTC 18

RESULT 9

LOCUS BG988031

DEFINITION MR2-HT1160-110101-005-a06 HT1160 Homo sapiens cDNA, mRNA sequence.

ACCESSION BG988031

VERSION BG988031.1 GI:14392101

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 450)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-HT1160-110101-005-a06&t3=2001-01-11&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 39
 High quality sequence stop: 325.

FEATURES

Location/Qualifiers
 1..450
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="Htl160"

/note="organ: head_neck; Vector: puc18; Site1: SmaI; Site2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 84 a 127 c 150 g 89 t

BASE COUNT
 ORIGIN

Alignment Scores:
 Pred. No.: 1.13 Length: 450
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.00% Indels: 0
 DB: 12 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x BG988031 (1-450)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 |||||
 Db 259 GACAAAGGCTGCCCTGCCGAGCAGAGCCAGCCCTCTGACGTCC 303

RESULT 10
 BM720186

LOCUS
 DEFINITION
 UI-E-E00-ahz-c-08-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone
 UI-E-E00-ahz-c-08-0-UI 5', mRNA sequence.

ACCESSION
 VERSION
 BM720186.1 GI:19039265

EST.

KEYWORDS

SOURCE

ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS
 TITLE
 Normalization and subtraction: two approaches to facilitate gene discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers
 1..466

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-E-E00-ahz-c-08-0-UI"

/tissue_type="fetal eye"

/dev_stage="fetal"

/lab_host="DHI0B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-E00"

/note="organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site1: EcoR I; Site2: Not I; UI-E-E00 is a cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 105 a 114 c 151 g 93 t 3 others
 ORIGIN

Alignment Scores:

Pred. No.: 1.18 Length: 466
 Score: 80.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.00% Indels: 0
 DB: 12 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x BM720186 (1-466)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 |||||

Db 18 GACAAAGGCTGCCCTGCCGAGCAGAGCCAGCCCTCTGACGTCC 62

RESULT 11

BM828479

LOCUS

DEFINITION

to Neu/ERBB-2 p185, mRNA sequence.

ACCESSION

VERSION

BM828479.1 GI:22860534

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 524)

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Moore, Kateri A.
 Department of Molecular Biology
 Princeton University
 217 Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544,
 USA

Tel: 609 258 0605

Fax: 609 258 2759

Email: kmoore@molbio.princeton.edu

These ESTs are derived from a subtracted cDNA library enriched for gene products expressed by a hematopoietic stem cell-supporting stromal cell line, AFT024.
 Seq primer: M13Reverse or T7.

FEATURES
source

Location/Qualifiers
1. .524
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/tissue_type="Fetal Liver"
/cell_type="Stromal cell"
/cell_line="AFT024"
/dev_stage="Embryonic day 14-14.5"
/lab_host="DH10s"
/clone_lib="AFT024-subtracted library"
/note="Organ: Fetal Liver; Vector: Sport 1; Site_1: Sal I; Site_2: Not I; Two directionally cloned cDNA libraries were made from a hematopoietic stem cell-supporting stromal cell line (AFT024) and from a non-supporting stromal cell line (2018). Subtractive hybridization was performed by hybridization of the target, AFT024, single stranded cDNA library in pSport1 to biotinylated RNA transcribed from the driver, 2018 cDNA library in pSport2 with inserts cloned in the complementary orientation. The AFT024-subtracted library contains 4.2x105 clones and is depleted of common housekeeping gene products eg. beta-actin and enriched for transcripts specific to AFT024. For detailed protocols and additional information please see our website at
http://stromalcell.princeton.edu."

BASE COUNT 110 a 149 c 157 g 107 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 1.37 Length: 524
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 13 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x BQ828479 (1-524)

Qy 16 GlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
|||||
Db 279 CAGAACGAGGACTTAGGCCCTCCAGCCCATGGACGACACCTTC 323

RESULT 12

CB719751
LOCUS CB719751 527 bp mRNA linear EST 10-APR-2003
DEFINITION AMGNNUC:NRDc1-00186-c7-A nrdg1 (10855) Rattus norvegicus cDNA clone nrdg1-00186-c7 5', mRNA sequence.

ACCESSION CB719751
VERSION CB719751.1 GI:29776893

KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 527)
AUTHORS Angen EST Program.
TITLE Angen Rat EST Program
JOURNAL Unpublished
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00186 row: c column: 7.

Location/Qualifiers

1. .527

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="nrdg1-00186-c7"

/tissue_type="Dorsal Root Ganglia"

FEATURES

source

Location/Qualifiers

1. .527

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="nrdg1-00186-c7"

/tissue_type="Dorsal Root Ganglia"

BASE COUNT 105 a 159 c 156 g 107 t
ORIGIN

Alignment Scores:

Pred. No.: 1.38 Length: 527
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x CB719751 (1-527)

Qy 16 GlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
|||||
Db 96 CAGAACGAGGACTTGGCCCATCCAGCCCATGGACGACGACCTTC 140

RESULT 13

BI194790
LOCUS BI194790 546 bp mRNA linear EST 10-JUL-2001
DEFINITION 602948150F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5091396 5', mRNA sequence.

ACCESSION BI194790

VERSION BI194790.1 GI:14649810

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 546)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCMI855 row: c column: 13

High quality sequence start: 10

High quality sequence stop: 495.

Location/Qualifiers

1. .546

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5091396"

/tissue_type="epithelioid carcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_42"

/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library. "

BASE COUNT 110 a 147 c 183 g 106 t

ORIGIN

Alignment Scores:

Pred. No.: 1.44 Length: 546

Score: 80.00 Matches: 15

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.00%	Indels:	0
DB:	12	Gaps:	0
JUNC_SEQ3_SEQ2RES991_ (1-30) x BI149790 (1-546)			
QY	1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15		
Db	223 GACAAGGGCTGCCCGCGAGCAGAGCCAGCCCTCTGACGTCC 267		
RESULT 14			
AA443351			
LOCUS	AA443351 547 bp mRNA linear EST 03-JUN-1997		
DEFINITION	zw85b05.r1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:783729 5' similar to gb:M11730 ERBB-2 RECEPTOR PROTEIN-TYROSINE KINASE PRECURSOR (HUMAN);, mRNA sequence.		
ACCESSION	AA443351		
VERSION	AA443351.1 GI:2156026		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 547)		
AUTHORS	Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,F., Lacy,M., Le.N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.		
TITLE	WashU-Merck EST Project 1997		
JOURNAL	Unpublished		
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -28ml3 rev2 Et from Amersham High quality sequence stop: 360.		
FEATURES	source		
Location/Qualifiers	1..547		
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="GDB:5981347"			
/db_xref="taxon:9606"			
/clone="IMAGE:783729"			
/dev_stage="8-9 weeks"			
/lab_host="DH10B"			
/clone_lib="Soares_total_fetus_Nb2HF8_9w"			
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGCGCTTAATTTTTTTTTTTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."			
BASE COUNT	115 a	151 c	172 g 108 t
ORIGIN	1 others		
Alignment Scores:			
Pred. No.:	1.44	Length:	547
Score:	80.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.00%	Indels:	0
DB:	9	Gaps:	0
JUNC_SEQ3_SEQ2RES991_ (1-30) x AA443351 (1-547)			

JUNC_SEQ3_SEQ2RES991_ (1-30) x BM720098 (1-561)

Qy	1	AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer	15
Db	400	GACAAGGGCTGCCCGCCGAGCAGAGAGAGCCGCCCTCTGACGTCC	444

Search completed: October 16, 2003, 17:03:59
Job time : 240.794 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 11:08:07 : Search time 31.0111 Seconds
(without alignments)
2540.503 Million cell updates/sec

JUNC_SEQ3_SEQ2RES991_
Perfect score: 160
Sequence: 1 DKGPAPQASPLTSQNEGLGPPMDSTF 30

Scoring table:
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1750203.seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/p/HOLLERMAN480/runat_15102003_131915_20662/app_query.fasta_1.4685
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=rpnb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-NAXLEN=2000000000 -USER=HOLLERMAN480 -CGN_1_1_397@runat_15102003_131915_20662
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -YGAPOP=10 -XGAPEXT=0.5
-YGAPOP=6 -YGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	80	50.0	201	13 US-10-109-213-3
				Sequence 3, Appli

2	80	50.0	3765	12	US-10-207-498-5	Sequence 5, Appli
3	80	50.0	3768	9	US-09-811-123-8	Sequence 8, Appli
4	80	50.0	3768	9	US-09-811-115-2	Sequence 2, Appli
5	80	50.0	3768	10	US-09-854-356-9	Sequence 9, Appli
6	80	50.0	3768	10	US-09-930-125-1	Sequence 1, Appli
7	80	50.0	3768	12	US-10-313-644-1	Sequence 11, Appli
8	80	50.0	3771	10	US-09-854-356-11	Sequence 11, Appli
9	80	50.0	3955	10	US-09-870-759-117	Sequence 117, App
10	80	50.0	3955	10	US-09-854-356-10	Sequence 10, Appli
11	80	50.0	3955	12	US-09-751-708A-117	Sequence 117, App
12	80	50.0	4473	11	US-09-441-411-5	Sequence 5, Appli
13	80	50.0	4473	12	US-10-101-510-81	Sequence 81, Appli
14	80	50.0	4473	14	US-10-146-473-32	Sequence 32, Appli
15	80	50.0	4473	14	US-10-207-655-44	Sequence 44, Appli
16	80	50.0	4530	10	US-09-877-177-11	Sequence 11, Appli
17	80	50.0	4530	12	US-10-007-926A-119	Sequence 119, App
18	80	50.0	4530	12	US-10-101-510-124	Sequence 124, App
19	80	50.0	4530	12	US-10-338-730-1	Sequence 1, Appli
20	80	50.0	4530	14	US-10-177-293-125	Sequence 125, App
21	80	50.0	4543	10	US-09-769-508-1	Sequence 1, Appli
22	80	50.0	4606	12	US-09-971-392-70	Sequence 70, Appli
23	80	50.0	4642	14	US-10-198-846-10896	Sequence 10896, A
24	80	50.0	9274	9	US-09-811-123-7	Sequence 7, Appli
25	80	50.0	9274	9	US-09-811-115-1	Sequence 1, Appli
26	74	46.2	1115	14	US-10-102-806-165	Sequence 165, App
27	74	46.2	1713	12	US-10-378-393-14	Sequence 14, Appli
28	74	46.2	1755	10	US-09-930-125-6	Sequence 6, Appli
29	74	46.2	1767	10	US-09-930-125-4	Sequence 4, Appli
30	74	46.2	1773	10	US-09-930-125-7	Sequence 7, Appli
31	74	46.2	1806	10	US-09-930-125-5	Sequence 5, Appli
32	74	46.2	2411	12	US-10-378-393-10	Sequence 10, Appli
33	65	40.6	23432	9	US-09-764-869-1332	Sequence 1332, Ap
34	65	40.6	23432	14	US-10-091-504-1332	Sequence 1332, Ap
35	60	37.5	111282	13	US-10-094-989-3	Sequence 3, Appli
36	59	36.9	4092	11	US-09-925-388-5	Sequence 5, Appli
37	59	36.9	4092	12	US-10-431-846-5	Sequence 5, Appli
38	58	36.2	1938	12	US-10-422-264-25	Sequence 25, Appli
39	58	36.2	2316	12	US-10-422-264-27	Sequence 27, Appli
40	58	36.2	2604	12	US-10-432-264-23	Sequence 23, Appli
41	58	36.2	3320	10	US-09-862-658-1	Sequence 1, Appli
42	58	36.2	3320	14	US-10-175-696-22	Sequence 22, Appli
43	58	36.2	3384	12	US-10-422-264-39	Sequence 29, Appli
44	57.5	35.9	363	11	US-09-918-995-28706	Sequence 28706, A
45	57.5	35.9	2691	10	US-09-967-768A-292	Sequence 292, App

ALIGNMENTS

RESULT 1

US-10-109-213-3

; Sequence 3, Application US/10109213

; Publication No. US20020168670A1

; GENERAL INFORMATION:

; APPLICANT: Ecker, David J.

; TITLE OF INVENTION: Identification Of Disease Predictive Nucleic Acids

; FILE REFERENCE: IBIS0009

; CURRENT APPLICATION NUMBER: US/10/109,213

; CURRENT FILING DATE: 2002-03-27

; PRIOR APPLICATION NUMBER: US/09/200,355

; PRIOR FILING DATE: 1998-11-25

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 201

; TYPE: RNA

; ORGANISM: Homo sapiens

US-10-109-213-3

Alignment Scores:

Pred. No.: 0.000792

Score: 80.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Length: 201

Matches: 15

Conservative: 0

Mismatches: 0


```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

```

```

Alignment Scores:
Pred. No.: 0.0194 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

```

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-854-356-9 (1-3768)

```

Qy 1 AsplysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGAGCTCC 1959

```

RESULT 6

```

US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930.125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-09-930-125-1

```

```

Alignment Scores:
Pred. No.: 0.0194 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

```

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-930-125-1 (1-3768)

```

Qy 1 AsplysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGAGCTCC 1959

```

RESULT 7

```

US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313.644
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-313-644-1

```

```

Alignment Scores:
Pred. No.: 0.0194 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 12 Gaps: 0

```

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-10-313-644-1 (1-3768)

```

Qy 1 AsplysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGAGCTCC 1959

```

RESULT 8

```

US-09-854-356-11
; Sequence 11, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854.356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 3771
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse HER-2/neu
US-09-854-356-11

```

```

Alignment Scores:
Pred. No.: 0.0195 Length: 3771
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0

```

```
;
; NAME/KEY: CDS
; LOCATION: (26)..(3799)
; OTHER INFORMATION: rat HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (26)..(1990)
; OTHER INFORMATION: extracellular domain (ECD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2057)..(3796)
; OTHER INFORMATION: intracellular domain (ICD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2188)..(3022)
; OTHER INFORMATION: kinase domain (KD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2999)..(3796)
; OTHER INFORMATION: phosphorylation domain (PD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2999)..(3173)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of rat HER-2/neu
US-09-854-356-10

Alignment Scores:
Pred. No.: 0.0205 Length: 3955
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-854-356-11 (1-3771)

QY 16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
|||||
Db 2971 CAGAACGAGGACTTAGGCCCTCCAGCCCATGGACGACGACCTTC 3015

RESULT 9
US-09-870-759-117
; Sequence 117, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 117
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17)..(3799)
; OTHER INFORMATION:
US-09-870-759-117

Alignment Scores:
Pred. No.: 0.0205 Length: 3955
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-870-759-117 (1-3955)

QY 16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
|||||
Db 2999 CAGAACGAGGACTTGGCCCATCCAGCCCATGGACGACGACCTTC 3043

RESULT 10
US-09-854-356-10
; Sequence 10, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: rat HER-2/neu cDNA
```

```
;
; NAME/KEY: CDS
; LOCATION: (26)..(3799)
; OTHER INFORMATION: rat HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (26)..(1990)
; OTHER INFORMATION: extracellular domain (ECD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2057)..(3796)
; OTHER INFORMATION: intracellular domain (ICD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2188)..(3022)
; OTHER INFORMATION: kinase domain (KD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2999)..(3796)
; OTHER INFORMATION: phosphorylation domain (PD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2999)..(3173)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of rat HER-2/neu
US-09-854-356-10

Alignment Scores:
Pred. No.: 0.0205 Length: 3955
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-854-356-10 (1-3955)

QY 16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
|||||
Db 2999 CAGAACGAGGACTTGGCCCATCCAGCCCATGGACGACGACCTTC 3043

RESULT 11
US-09-751-708A-117
; Sequence 117, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 117
; LENGTH: 3955
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17)..(3799)
; OTHER INFORMATION:
US-09-751-708A-117

Alignment Scores:
Pred. No.: 0.0205 Length: 3955
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-751-708A-117 (1-3955)

QY 16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
|||||
Db 2999 CAGAACGAGGACTTGGCCCATCCAGCCCATGGACGACGACCTTC 3043
```

RESULT 12

US-09-441-411-5
; Sequence 5, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-5

Alignment Scores:
Pred. No.: 0.0234 Length: 4473
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 11 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-441-411-5 (1-4473)

Qy 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGCTGCCCGCGCAGAGAGAGCCAGCCCTCTGACGTCC 2133

RESULT 13

US-10-101-510-81
; Sequence 81, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-81

Alignment Scores:
Pred. No.: 0.0234 Length: 4473
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-10-101-510-81 (1-4473)

Qy 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGCTGCCCGCGCAGAGAGAGCCAGCCCTCTGACGTCC 2133

RESULT 14

US-10-146-473-32
; Sequence 32, Application US/10146473

; Publication No. US20030108888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-473-32

Alignment Scores:
Pred. No.: 0.0234 Length: 4473
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 14 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-10-146-473-32 (1-4473)

Qy 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGCTGCCCGCGCAGAGAGAGCCAGCCCTCTGACGTCC 2133

RESULT 15

US-10-207-655-44
; Sequence 44, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069,401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-207-655-44

Alignment Scores:
Pred. No.: 0.0234 Length: 4473
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 14 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-10-207-655-44 (1-4473)

Qy 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2089 GACAAGGCTGCCCGCGCAGAGAGAGCCAGCCCTCTGACGTCC 2133

Search completed: October 17, 2003, 03:52:37
Job time : 36.0111 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 28.4941 Seconds
(without alignments)
2842.104 Million cell updates/sec

Title: JUNC_SEQ3_SEQ2RES991_
Perfect score: 160
Sequence: 1 DKGCPAEQASPLTSQNEDLGPPSPMDSTF 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/Cyn2.1/USPFO_spool_p/HOLLERAN480/runat_15102003_131912_20526/app_query.fasta_1.4685
-DB=N.Geneseq_19Jun03 -QMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTEXT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480_@CGN_1_1_1758_@runat_15102003_131912_20526 -NCPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	150	93.8	2763	24	ABA92252	Mouse Her-2/neu ex
2	150	93.8	2781	24	ABA92253	Her-2/neu extracel
3	82	51.2	2385	18	AAAT72725	Her2-GM-CSF immuno
4	80	50.0	69	21	AAA14815	DNA encoding a wil
5	80	50.0	201	21	AAAG2125	Her2 normal form t
6	80	50.0	2125	19	AAV21727	Humanised vector p
7	80	50.0	2871	21	AAZ50586	DC8scFv-erbB2EC fu
8	80	50.0	3600	21	AAAS9736	Human HER-2/neu co
9	80	50.0	3678	24	ABK86207	cDNA encoding huma
10	80	50.0	3768	17	AAAT07339	HER-2/neu oncogene
11	80	50.0	3768	20	AAAT01912	Human HER-2/neu on
12	80	50.0	3768	21	AAA09455	Human heregulin 2
13	80	50.0	3768	22	AAH23392	Human HER-2/neu pr
14	80	50.0	3768	24	ABZ35744	Human ERBB2 polynu
15	80	50.0	3768	24	ABX09987	Human ERBB2 DNA fr
16	80	50.0	3768	24	AAAD3935	Human HER-2 CDNA.
17	80	50.0	3768	24	AAAD3986	Human ERBB2 DNA SE
18	80	50.0	3768	24	ABV78168	Human Her-2/neu pr
19	80	50.0	3768	24	AAAD32743	Human Her-2/neu CD
20	80	50.0	3768	24	ABA92250	Human Her-2/neu DN
21	80	50.0	3768	24	ABK10730	Human polynucleoti
22	80	50.0	3768	24	ABK14058	Human HER2 (ERBB2)
23	80	50.0	3768	24	ABK14058	Mouse Her-2/neu CD
24	80	50.0	3771	21	AAA89737	Nucleotide sequenc
25	80	50.0	3771	22	AAH42210	Mouse Her-2/neu CD
26	80	50.0	3771	24	ABA92251	Rat neu promoter.
27	80	50.0	3955	16	AAAT01590	Rat HER-2/neu prot
28	80	50.0	3955	21	AAA89753	Sequence encoding
29	80	50.0	4299	14	AAQA6083	cDNA encoding the
30	80	50.0	4472	21	AAA14812	Human tumour anti
31	80	50.0	4473	19	AB076220	HER-2 nucleic acid
32	80	50.0	4473	20	AAZ1071	Human gene express
33	80	50.0	4473	24	ABZ34969	Human Her-2 DNA.
34	80	50.0	4473	16	AAAT01585	Her-2/neu (ERBB2/c
35	80	50.0	4530	18	AAAT71253	Human HER2 gene.
36	80	50.0	4530	21	AAZ60815	Nucleotide sequenc
37	80	50.0	4530	22	AAAD19731	Human tyrosine kin
38	80	50.0	4530	24	ABZ35012	Human gene express
39	80	50.0	4530	24	ABV94128	Breast carcinoma r
40	80	50.0	4530	24	ABN85585	Human HER2-neu SEQ
41	80	50.0	4530	24	ABK83918	Human cDNA differe
42	80	50.0	4530	25	ACC50139	Breast cancer asso
43	80	50.0	4530	25	ABQ83856	Human Her2/Neu enc
44	80	50.0	9274	22	AAF24237	HER2 transgene pla

ALIGNMENTS

RESULT 1
ABA92252
ID ABA92252 standard; cDNA; 2763 BP.
AC ABA92252;
XX
XX 17-JUN-2002 (first entry)
DT
XX Mouse Her-2/neu extracellular-phosphorylation domain fusion cDNA.
DE
XX Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
KW receptor; mouse; gene therapy; gene; ss.
XX
XX Mus musculus.
OS
XX
XX Key Location/Qualifiers
FH 1..2763
FT CDS


```
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1918 GAACGAGCTGCCACGACAGACAGAGCCAGCCAGTGCAGCTCTCAGAACGAGGACTTA 1977
QY 21 GlyProSerSerProMetAspSerThrPhe 30
Db 1978 GGCCCTCCAGCCCATGGACAGCACCTTC 2007

RESULT 3
AAAT72725
ID AAAT72725 standard; cDNA; 2385 BP.
AC AAAT72725;
XX
XX
XX 17-SEP-1997 (first entry)
XX
DE Her2-GM-CSF immunostimulant fusion protein DNA.
XX
KW Her2-GM-CSF; granulocyte macrophage colony stimulating factor;
KW growth factor receptor; oncogene; immunostimulant; cancer;
KW therapy; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 11..2359
XX /tag= a
XX /product= GM-CSF-Her2 fusion protein
XX mRNA 11..1969
XX /tag= b
XX /product= Her2
XX mRNA 1970..1975
XX /tag= c
XX /product= Leu-Glu linker
XX mRNA 1976..2359
XX /tag= d
XX /product= GM-CSF
XX
XX WO9724438-A1.
XX
XX 10-JUL-1997.
XX
XX 23-DEC-1996; 96WO-US20241.
XX
XX 28-DEC-1995; 95US-0579823.
XX
XX (ACT1-) ACTIVATED CELL THERAPY INC.
XX
XX Laus R, Ruegg CL, Wu H;
XX
XX WPI; 1997-363674/33.
XX P-PSDB; AAW19764.
XX
XX Potent APC that activates T-cells to give multivalent cellular
XX immune response - can also induce a cytotoxic T-cell response in a
XX vertebrate subject
XX
XX Disclosure; Fig 8; 45pp; English.
XX
XX A nucleic acid molecule (AAAT72725) codes for a fusion protein
XX (AAW19764) comprising granulocyte-macrophage colony stimulating
XX factor (GM-CSF) and Her2, a growth factor receptor that is
XX over-expressed in breast and ovarian cancer cells. It was
XX prepared by PCR amplification of Her2 cDNA from a breast cancer
XX cell line and fusion to GM-CSF cDNA. Fusion expression vectors can
XX be used to transfect mammalian and insect cells. The Her2-GM-CSF
XX fusion protein is used to generate anti-Her2 immunity. Tumour
XX cells are eliminated by cytotoxic T lymphocytes activated in vivo
XX or in vitro by exposure to antigen-presenting cells exposed to the
XX fusion protein.
XX
XX Sequence 2385 BP; 488 A; 780 C; 677 G; 440 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 0.0564 Length: 2385
Score: 82.00 Matches: 17
Percent Similarity: 68.00% Conservative: 0
Best Local Similarity: 68.00% Mismatches: 8
Query Match: 51.25% Indels: 0
DB: 18 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x AAT72725 (1-2385)
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGCTGCCCGCCGACAGAGAGAGCCAGCCCTCTGACGTCCCTCGAGGCACCCGCC 1984
QY 21 GlyProSerSerPro 25
Db 1985 CGCTCGCCAGCCCC 1999

RESULT 4
AAAT4815
ID AAAT4815 standard; cDNA; 69 BP.
XX
XX AC AAAT4815;
XX
XX 08-AUG-2000 (first entry)
XX
XX DNA encoding a wild type erbB-2 receptor protein fragment.
XX
XX SPLICE erbB-2 receptor protein; cell transformation disorder; cancer;
XX tumor cell proliferation; tissue degeneration; arthropathy;
XX bone resorption; inflammatory disease; degenerative disorder;
XX wound healing; ss.
XX
XX Homo sapiens.
XX
XX WO200020579-A1.
XX
XX 13-APR-2000.
XX
XX 01-OCT-1999; 99WO-CA00912.
XX
XX 02-OCT-1998; 98US-0165192.
XX
XX (UYMC-) UNIV MCMASTER.
XX
XX Muller WJ, Siegel PM;
XX
XX WPI; 2000-303768/26.
XX P-PSDB; AAY84681.
XX
XX Nucleic acid encoding an erbB 2 receptor protein designated SPLICE
XX erbB-2, inhibitors of the protein are useful for treatment of cancer -
XX Example 2; Fig 3B; 60pp; English.
XX
XX The present sequence encodes a peptide of wild type erbB-2 receptor
XX protein, where the splicing event occurs. An alternatively spliced
XX variant of erbB-2 receptor protein, designated SPLICE erbB-2 receptor
XX protein, has an in-frame deletion of 16 amino acids, 2 of which are
XX conserved cysteine residues, compared to the unspliced protein. The
XX erbB-2 polynucleotide is used to construct probes for detecting
XX disorders of cell transformation such as cancer. Antibodies to the
XX protein may be used to detect SPLICE erbB-2 in a sample. Agents
XX (e.g. antisense oligonucleotides) which inhibit the expression of
XX SPLICE erbB-2 are useful for reducing tumor cell proliferation and
XX treating cancer. Substances which stimulate SPLICE erbB-2 are useful
XX for treating conditions involving damaged cells including conditions in
XX which degeneration of tissue occurs, such as arthropathy, bone
XX resorption, inflammatory diseases, degenerative disorders of the
XX central nervous system and wound healing.
XX
XX Sequence 69 BP; 13 A; 26 C; 20 G; 10 T; 0 other;
```


PN US5869445-A.
XX
PD 09-FEB-1999.
XX
XX 01-APR-1996; 96US-0625101.
XX
XX 01-APR-1996; 96US-0625101.
PR 17-MAR-1993; 93US-0033644.
PR 12-AUG-1993; 93US-0106112.
PR 31-MAR-1995; 95US-0414417.
XX
PA (UNIW) UNIV WASHINGTON.
XX
XX Cheever MA, Disis ML;
PI WPI; 1999-152835/13.
DR P-PSDB; AAW92406.
XX
XX Use of HER-2/neu polypeptides - for eliciting an immune response to
PT an HER-2/neu associated malignancy, particularly for treating or
PT preventing tumours
XX
XX Claim 1a; Column 23-32; 26pp; English.
XX
XX This sequence encodes the human HER-2/neu oncogene protein. A fragment
CC of this protein is used in a method for eliciting or enhancing an immune
CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
CC B cells to produce an immune response to the HER-2/neu protein. The
CC method can be used for immunisation against a malignancy in which the
CC HER-2/neu oncogene is associated and in the treatment of an existing
CC tumour, or to prevent tumour occurrence or reoccurrence.
XX
SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.192 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 20 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x AAX01912 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAAGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTC 1959

RESULT 12
AAA09455
ID AAA09455 standard; DNA; 3768 BP.
XX
AC AAA09455;
XX
XX 10-AUG-2000 (first entry)
DE Human heregulin 2 (Her2) coding sequence.
XX
XX Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
KW self-protein; cell-associated peptide antigen; foreign epitope;
KW cancer; breast cancer; prostate cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
PR 20-OCT-1998; 98US-0105011.
XX

PA (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
DR P-PSDB; AAY92620.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
XX Claim 62; Page 187-193; 220pp; English.
XX
XX The claims detail a method for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, human
CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
XX PSM, FGF8b and Her2, respectively.
XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.192 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x AAA09455 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAAGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTC 1959

RESULT 13
AAH23392
ID AAH23392 standard; DNA; 3768 BP.
XX
XX AAH23392;
AC
XX
XX 25-SEP-2001 (first entry)
DT
XX
XX Human HER-2/neu protein encoding DNA.
DE
XX
XX Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1..3768
FT CDS /*tag= a
FT /*product= "HER-2/neu protein"
XX
XX WO200153463-A2.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 19-JAN-2001; 2001WO-US01850.
PF
XX
XX 21-JAN-2000; 2000US-0177545.
PR
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 432.026 Seconds
(without alignments)
2840.777 Million cell updates/sec

Title: JUNC_SEQ3_SEQ2RES991_
Perfect score: 160
Sequence: 1 DKGCPAEORASPLTSQNEIDLGSPSPMDSTF 30

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/csn2_1/USPTO_spool_P/HOLLERAN480/runat_15102003_131912_20536/app_query.fasta_1.4685
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480_SCGN_1_1_22724_@runat_15102003_131912_20536 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb.ba: *
2: gb.htg: *
3: gb.in: *
4: gb.om: *
5: gb.ov: *
6: gb.pat: *
7: gb.ph: *
8: gb.pl: *
9: gb.pr: *
10: gb.ro: *
11: gb.sts: *
12: gb.sy: *
13: gb.un: *
14: gb.vi: *
15: em.ba: *
16: em.fun: *
17: em.hum: *
18: em.in: *
19: em.mu: *
20: em.om: *
21: em.or: *
22: em.ov: *
23: em.pat: *
24: em.ph: *
25: em.pl: *
26: em.ro: *
27: em.sts: *
28: em.un: *

29: em.vi: *
30: em.htg_hum: *
31: em.htg_inv: *
32: em.htg_other: *
33: em.htg_mus: *
34: em.htg_pln: *
35: em.htg_rod: *
36: em.htg_mam: *
37: em.htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	150	93.8	2763	6	AX380942 Sequence
2	150	93.8	2781	6	AX380944 Sequence
C 3	91.5	57.2	186134	10	AL591390 Mouse DNA
C 4	91.5	57.2	214019	2	AC064803 Mus muscu
5	82	51.2	2385	6	AR082744 Sequence
6	82	51.2	2385	6	AR099963 Sequence
7	82	51.2	2385	6	AR143949 Sequence
8	80	50.0	153	6	AX022984 Sequence
9	80	50.0	201	6	AR229723 Sequence
10	80	50.0	2083	10	BC027080 Mus muscu
11	80	50.0	2871	6	AX023363 Sequence
12	80	50.0	3678	6	AX050514 Sequence
13	80	50.0	3768	6	AR034479 Sequence
14	80	50.0	3768	6	AX060704 Sequence
15	80	50.0	3768	6	AX201817 Sequence
16	80	50.0	3768	6	AX380923 Sequence
17	80	50.0	3768	6	AX384604 Sequence
18	80	50.0	3768	6	AX465456 Sequence
19	80	50.0	3768	6	AX467229 Sequence
20	80	50.0	3768	6	AX481438 Sequence
21	80	50.0	3771	6	AX189662 Sequence
22	80	50.0	3771	6	AX380925 Sequence
23	80	50.0	3955	6	AX380924 Sequence
24	80	50.0	3955	6	I21129 Sequence 14
25	80	50.0	3955	6	I59750 Sequence 14
26	80	50.0	3955	10	RNNEUR X03362 Rat mRNA fo
27	80	50.0	4473	6	AR080259 Sequence
28	80	50.0	4473	6	AR167390 Sequence
29	80	50.0	4473	9	HSERB2R X03363 Human c-erb
30	80	50.0	4530	6	AR202597 Sequence
31	80	50.0	4530	6	AR283481 Sequence
32	80	50.0	4530	6	AX282577 Sequence
33	80	50.0	4530	6	AX587649 Sequence
34	80	50.0	4530	6	AX644071 Sequence
35	80	50.0	4530	6	BD005474 Cellular
36	80	50.0	4530	6	I21124 Sequence 9
37	80	50.0	4530	6	I59745 Sequence 9
38	80	50.0	4530	9	HUMHER2A M11730 Human tyros
39	80	50.0	4694	10	BC046811 Mus muscu
40	80	50.0	4695	10	BC053078 Mus muscu
41	80	50.0	4727	10	AY116182 Rattus no
42	80	50.0	9274	6	AX060703 Sequence
43	80	50.0	13450	9	AB096612 Homo sapi
44	80	50.0	16572	9	AB096614 Homo sapi
45	80	50.0	20271	9	AB096613 Homo sapi

ALIGNMENTS

RESULT 1

JUNC_SEQ3_SEQ2RES991_ (1-30) x AX380944 (1-2781)

Qy 1 AsplyGlyCysProAlaGluGlnAinrAlaSerProLeuThrSerGlnAsnGluAspLeu 20
:::.....|:::.....|:::.....|:::.....|:::.....|:::.....|

Db 1918 GAACGAGGTGGCCAGCAGACAGACAGAGCCAGGCCCTGCAGTCTCAGAACGAGACTTA 1977

Qy 21 GlyProSerSerProMetAspSerThrPhe 30
|||.....|.....|.....|.....|.....|.....|

Db 1978 GGCCCCCTCCAGCCCCCATGGACAGCACCTTC 2007

RESULT 3

AL591390/c AL591390 186134 bp DNA linear ROD 17-NOV-2001

LOCUS Mouse DNA sequence from clone RP23-355L10 on chromosome 11,
DEFINITION complete sequence.

ACCESSION AL591390

VERSION AL591390.8 GI:17017767

KEYWORDS HTG.

SOURCE Mus musculus

ORGANISM Mus musculus (house mouse)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 186134)
Phillimore,B.
Direct Submission
Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 20, 2001 this sequence version replaced gi:16535512.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-355L10 is
from the RPI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone
RP23-355L10 It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP23-355L10 is at 136134 in this
sequence. The true right end of clone RP23-438D7 is at 2000 in this
sequence.

FEATURES

source Location/Qualifiers
1..186134
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="11"
/clone="RP23-355L10"
/clone_lib="RPCI-23"
17991..18044
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."

misc_feature misc_feature
47794..47812
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."
46287 a 47900 c 47689 g 44258 t

BASE COUNT 46287 a 47900 c 47689 g 44258 t

ORIGIN

misc_feature

misc_feature 10885..122875
 /note="assembly_fragment"
 misc_feature 122976..139026
 /note="assembly_fragment"
 misc_feature 139127..159729
 /note="assembly_fragment"
 misc_feature 159830..178173
 /note="assembly_fragment"
 misc_feature 178274..200970
 /note="assembly_fragment"
 misc_feature 201071..214019
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right"
 BASE COUNT 52809 a 53348 c 53260 g 53244 t 1358 others.
 ORIGIN

Alignment Scores:
 Pred. No.: 0.758 Length: 214019
 Score: 91.50 Matches: 20
 Percent Similarity: 72.41% Conservative: 1
 Best Local Similarity: 68.97% Mismatches: 5
 Query Match: 57.19% Indels: 3
 DB: 2 Gaps: 1

JUNC_SEQ3_SEQ2RES991_ (1-30) x AC064803 (1-214019)

QY 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21
 Db 53608 GAGGGCTGT-----CACCTCGTGTCTCCCTTCCTTAGAACGAGGACTTAGGC 53558
 QY 22 ProSerSerProMetAspSerThrPhe 30
 Db 53557 CCCTCCAGCCCATGACACCACTTC 53531

RESULT 5
 AR082744
 LOCUS AR082744 2385 bp DNA linear PAT 01-SEP-2000
 DEFINITION Sequence 3 from patent US 5976546.
 ACCESSION AR082744
 VERSION AR082744.1 GI:10009534
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2385)
 AUTHORS Laus,R., Ruegg,C.Landon. and Wu,H.
 TITLE Immunostimulatory compositions
 JOURNAL Patent: US 5976546-A 3 02-NOV-1999;
 FEATURES Location/Qualifiers
 source 1..2385
 /organism="unknown"

BASE COUNT 488 a 780 c 677 g 440 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.165 Length: 2385
 Score: 82.00 Matches: 17
 Percent Similarity: 68.00% Conservative: 0
 Best Local Similarity: 68.00% Mismatches: 8
 Query Match: 51.25% Indels: 0
 DB: 6 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x AR082744 (1-2385)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
 Db 1925 GACAAGGGCTGCCCGCGAGCAGAGCCAGCCCTCTGACGTCCTCGAGGCACCCGCC 1984
 QY 21 GlyProSerSerPro 25
 Db 1985 CGCTCGCCAGGCC 1999

RESULT 6
 AR099963
 LOCUS AR099963 2385 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 3 from patent US 6080409.
 ACCESSION AR099963
 VERSION AR099963.1 GI:12810411
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2385)
 AUTHORS Laus,R., Ruegg,C.Landon. and Wu,H.
 TITLE Immunostimulatory method
 JOURNAL Patent: US 6080409-A 3 27-JUN-2000;
 FEATURES Location/Qualifiers
 source 1..2385
 /organism="unknown"

BASE COUNT 488 a 780 c 677 g 440 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.165 Length: 2385
 Score: 82.00 Matches: 17
 Percent Similarity: 68.00% Conservative: 0
 Best Local Similarity: 68.00% Mismatches: 8
 Query Match: 51.25% Indels: 0
 DB: 6 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x AR099963 (1-2385)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
 Db 1925 GACAAGGGCTGCCCGCGAGCAGAGCCAGCCCTCTGACGTCCTCGAGGCACCCGCC 1984
 QY 21 GlyProSerSerPro 25
 Db 1985 CGCTCGCCAGGCC 1999

RESULT 7
 AR143949
 LOCUS AR143949 2385 bp DNA linear PAT 08-AUG-2001
 DEFINITION Sequence 3 from patent US 6210662.
 ACCESSION AR143949
 VERSION AR143949.1 GI:15105816
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2385)
 AUTHORS Laus,R., Ruegg,C.Landon. and Wu,H.
 TITLE Immunostimulatory composition
 JOURNAL Patent: US 6210662-A 3 03-APR-2001;
 FEATURES Location/Qualifiers
 source 1..2385
 /organism="unknown"

BASE COUNT 488 a 780 c 677 g 440 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.165 Length: 2385
 Score: 82.00 Matches: 17
 Percent Similarity: 68.00% Conservative: 0
 Best Local Similarity: 68.00% Mismatches: 8
 Query Match: 51.25% Indels: 0
 DB: 6 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x AR143949 (1-2385)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
 Db 1925 GACAAGGGCTGCCCGCGAGCAGAGCCAGCCCTCTGACGTCCTCGAGGCACCCGCC 1984
 QY 21 GlyProSerSerPro 25

```

Db      1985 CGTCGCCAGCCGCC 1999
|||||
RESULT 8
LOCUS   AX022984
DEFINITION Sequence 4 from Patent EP0919622.
ACCESSION AX022984
VERSION   AX022984.1 GI:10046479
KEYWORDS
SOURCE   unidentified
ORGANISM unidentified
REFERENCE
1 Springer,C.J. and Marais,R.
TITLE Surface expression of enzyme in gene directed prodrgug therapy
JOURNAL Patent: EP 0919622-A 4 02-JUN-1999;
CANCER RES CAMPAIGN TECH (CB)
FEATURES
source
1. 153
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 27 a 43 c 51 g 32 t
ORIGIN

Alignment Scores:
Pred. No.: 0.0198 Length: 153
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
Gaps: 0
DB: 6

JUNC_SEQ3_SEQ2RES991_ (1-30) x AX022984 (1-153)
Qy 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 10 GACAAGGCTGCCCCCGCGGACGACGAGCCAGCCCTCTGAGCTCC 54

RESULT 9
LOCUS   AR229723
DEFINITION Sequence 3 from patent US 6451524.
ACCESSION AR229723
VERSION   AR229723.1 GI:27269547
KEYWORDS
SOURCE   Unknown.
ORGANISM Unclassified.
REFERENCE
1 (bases 1 to 201)
AUTHORS Ecker,D.J.
TITLE Identification of disease predictive nucleic acids
JOURNAL Patent: US 6451524-A 3 17-SEP-2002;
FEATURES
source
1. 201
/organism="unknown"
BASE COUNT 34 a 67 c 56 g 44 t
ORIGIN

Alignment Scores:
Pred. No.: 0.026 Length: 201
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
Gaps: 0
DB: 6

JUNC_SEQ3_SEQ2RES991_ (1-30) x AR229723 (1-201)
Qy 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 116 GACAAGGCTGCCCCCGCGGACGACGAGCCAGCCCTCTGAGCTCC 160

```

RESULT 10
LOCUS
DEFINITION

BC027080 2083 bp mRNA linear ROD 16-APR-2003
Mus musculus v-erb-b2 erythroblastic leukemia viral oncogene
homolog 2, neuro/glioblastoma derived oncogene homolog (avian),
mRNA (CDNA clone MGC:38648 IMAGE:5356166), complete cds.

BC027080
BC027080.1 GI:20071980

ACCESSION
VERSION
KEYWORDS

SOURCE
ORGANISM

Mus musculus (house mouse)
MGC.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2083)

REFERENCE

AUTHORS

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalski,U., Smalius,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

cDNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 54 Row: c Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

Location/Qualifiers.

1. .2083

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:38648 IMAGE:5356166"

FEATURES

source

```

/tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI_CGAP_Mam6"
/lab_host="DH103"
/notes=Vector: pCMV-SPORT6"
1. .2083
/genes="Erbb2"
/notes="synonyms: Neu, HER-2, c-erbB2, HER2, c-neu"
/db_xref="LocusID:13866"
/db_xref="MGI:95410"
459. .1562
/codon_start=1
/product="Erbb2 protein"
/protein_id="AAH27080.1"
/db_xref="GI:20071981"
/db_xref="LocusID:13866"
/translation="MALESILRRRFTHQSDVMSYGVTVVWELMTFGAKPDGIPAREIP
DLLEKERLPQPICITDYMIMVKWMDISECRPRELSEFSRMARDPQRFVVIQ
NEDLGSSPMDSFYRSLLEDMDGELVDAEYLVPOQGFSPDPALGTGSTARHRH
SSSARGSGGELTGLPESEEPSPRLAPSEGAGSDVFDGLAVGTGQLSLSPHDL
SPLOYSDEPTLPLPETDGYVAPLACSPQEVYNQPEVRPOSPLTPEGPPPIRAG
ATLERPKTILSPCKNGVKDVFAGGAVENPEYLA PRAGTASQPHSPAPSPADNLYY
WQNSSEQPPSTFTGTTAENPEYLGIDVPV"
471 a 588 c 574 g 450 t
BASE COUNT
ORIGIN

Alignment Scores:
Pred. No.: 0.27 Length: 2083
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x BC027080 (1-2083)

QY 16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
|||||
Db 762 CAGAACGAGGACTTAGGCCCTCCAGCCCATGACAGCACCTC 806

RESULT 11
AX023363
LOCUS 2871 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 34 from Patent WO0006605.
ACCESSION AX023363
VERSION AX023363.1 GI:10183775
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Kufer,P., Zettl,F., Dreier,T., Baeuerle,P.A. and Borschert,K.
Heteromabinobodies
Patent: WO 0006605-A 34 10-FEB-2000;
KUFER PETER (DE) ; ZETTL FLORIAN (DE) ; DREIER TORSTEN (DE) ;
BAEUERLE PATRICK A (DE) ; BORSCHERT KATRIN (DE) ; MICROMET GES FUER
BIOMEDIZINIS (DE)
Location/Qualifiers
1. .2871
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
10. .2865
/notes="unnamed protein product"
/codon_start=1
/db_xref="GI:10183776"
/translation="MGWSCILFLVATATGVSHSDIOLTSOPATMSAPGKVTMTCTCA
SSVSYSLLHWQKSGAPKLIYXTSNLASGVPARFSGSGTSYSLTISVVEAD
AIIYICQISGYFTTGGTKLEIGSGGGGGGGGGGQVRLQSGAVKLPFGASV
KLSCASGYTLTSYLLHWQKWPGRGLEWIGRIDPNSGGTKYDKFKSKATLTVDKPS
STAYMLSSLTSDSAVYICARWDYWGQGTTVTVSSVSTPLGDTTHTSGKPLDGEYFTL

```

```

QIRGRERFERNEALELKDAQAGKEPGGGGGSTQVCTGTMKRLRASPETHLI
DMLRLHYOCQVVOGNLETYLPTNASLQDIOBVOGVYLIHAHQVQVPLQRLRI
VRGTOLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLTLILKGVLIQRP
QICYQDTILWKDIFHKNNQLALTLIDTNSRACHPCSPCKGSRGSESDQSLTR
TVAGGACRCKGPLPTDCCHEQCAAGCTGKHSDCLACLNHSHGICELRCPALVTYN
TDTFESMPNEGRYTFEGASCVTACTPYNLSTDVSGCTLYCELHNQSVTADGTQCEK
CSKPCARVCYGLGMEHLREYRAVTSANIQEFAGCKIFGSLAFIPESFDGDPASNTAP
LQPEQLQVFTLEITGYLYISAWPSLPLDSVFQNLQVIRGRLHNGAYSLTLOGIG
ISWLGRLSRELGSGLALJHNTHLQFVTPMDQLFRNPQALHLHTANRPEDECVGE
GLACHQLCARGHCWGPGTQCVCNCSQFLRGQECVEBCRVLQGLPRIYVNRHCLPCHP
ECQPQNSVTFCFGEADQVACAHYKPPFCVARCPGKGVKPDLSYMPIWKFPDEEGAC
QPCPINCHTSCVDLDDKQCPAEQRASPLTSGHHHHH"
598 a 868 c 834 g 571 t
BASE COUNT
ORIGIN

Alignment Scores:
Pred. No.: 0.372 Length: 2871
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x AX023363 (1-2871)

QY 1 AspLYSGLYCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 2797 GACAAGGGCTGCCGCCGAGCAGAGCAGCCCTCTGACGTCC 2841

RESULT 12
LOCUS 3678 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 7 from Patent WO0240059.
ACCESSION AX505114
VERSION AX505114.1 GI:23386421
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Mincheff,M.S., Loukinov,D.I. and Zoubak,S.
Methods and compositions for inducing cell-mediated immune
responses
Patent: WO 0240059-A 7 23-MAY-2002;
American Foundation for Biological Research Inc. (US) ; Mincheff,
Milcho S. (US) ; Loukinov, Dmitri I. (US) ; Zoubak, Serguei (US)
Location/Qualifiers
1. .3678
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
7. .3678
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD48804.1"
/db_xref="GI:23386422"
/translation="MRLPASPEPETHLMDLRLHYOCQVVOGNLELTLYPTNASLSFLQD
IQEVQVLIHAHQVQVPLQRLRIYRGQLFEDNTALAVLDNGDPLNNTPTVTGASP
GGLEQLRSLTEILKGVLIQRPOLCYQDTILWKDIFHKNNQLALTLIDTNRSRAC
HPCSPCKGSRGSESDQSLTRTVAGGACRCKGPIPTDCCHEQCAAGCTGPKHS
DCLACLHFNHSGICELRCPALVTYNSTDVSGCTLYCELHNQSVTADGTQCEK
GCTLYVCPLNHQEVTADGTQCEKSKPCARVCYGLGMEHLREYRAVTSANIQEFAG
CKKIFGSLAFIPESFDGDPASNTAPLQPEQLQVFTLEITGYLYISAWPSLPLDSV
FQNLQVIRGRLHNGAYSLTLOGIGISWLGRLSRELGSGLALJHNTHLQFVTPMD
DQLFRNPQALHLHTANRPEDECVGEGLACHQLCARGHCWGPGTQCVCNCSQFLRGQEC
VEBCRVLQGLPRIYVNRHCLPCHPQNSVTFCFGEADQVACAHYKPPFCVARCPGK
RCPSGVGPDLSPYPIWKFPDEEGACQPCINCTSCVDLDDKQCPAEQRASPLTSGH
AVVGIILVYVGVVGLILKRRQOKIRKTYMRELLQETELVEPLTSPGAMPNQAOMRI
LKETELRKVYLGSGAGFYVYKGIWIPDGENYKIPVAKIVLRENTSPKANKEILDEAY
VMAGVSGPYVSRLLIGICLTSTVOLTPYGLDHLHVRNREGRSGSLDNLNMCQIA
KMSYLEDVRLRFTHQSDVMSYGVTVVWELMTFGAKPDGIPAREIPDLLEKERLPQPI
MALESILRRRFTHQSDVMSYGVTVVWELMTFGAKPDGIPAREIPDLLEKERLPQPI

```


ORIGIN

Alignment Scores: Length: 3768
 Pred. No.: 0.488 Matches: 15
 Score: 80.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 50.00% Gaps: 0
 DB: 6

 JUNC_SEQ3_SEQ2RES991_ (1-30) x AX201817 (1-3768)
 QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
 Db 1915 GACAAGGGCTGCCCCCGCGAGAGAGAGCCCTCTGACGTCC 1959

Search completed: October 16, 2003, 10:58:43
 Job time : 468.208 secs